◆基礎数理講座◆

3

線形代数汎論

◆

伊理　正夫
［著］

朝倉書店

まえがき

現在我が国では(海外でも)「行列」,「行列式」,「線形代数」,等の言葉を本の表題あるいは一部の章の表題に含む本は数え切れないほどある．そこに新たに一つを付け加えることにどれだけの意味があるか，疑問に思われる方も少なくないかもしれない．しかし，私には，現存のそれらの本はどれも大同小異にみえる．やや暴言をお許し頂ければ，怠惰な学生と広範囲の応用の経験に乏しいのに応用系の学生を教えなければならない教師とに阿(おもね)るかのごとくに書かれた“分かりやすくて通り一遍の”教科書か，著者が“自分がどこまで抽象的にしかも厳密に理解しているかをひけらかす”ような数学者のための数学専門書かのどちらかで，数学の利用者，消費者をほとんど無視したものばかりのようである．私は，長年多くの応用分野で線形代数に関連した方法を利用し，また不足しているところは必要に応じて自前で補ったりしながら数理工学的な研究を続けてきた者であるが，現在遍在している上記のような本の著者達とは思い切って立場を変えて，利用者の観点からおよそ何かの役に立ちそうなものを体系的に整理して一冊の本に纏めてみるのも無駄ではなかろうと常々考えていた．

そのようなとき，岩波書店から岩波講座「応用数学」の刊行の企画があり，私がその中の 2 分冊「線形代数 I，II」を担当することになった．それは 1993 年，1994 年に刊行され，その後局所的な改訂を経て 1997 年に第 2 次刊行が，そしてそれら 2 分冊を 1 冊に纏めた単行本「一般線形代数」が同書店から 2003 年に刊行されたが，技術的な制約から，内容の選択の改善や誤謬の訂正など，根本的な改訂は出来なかった．多くの読者から従来の数学書とは異なる応用志向の書物であるとの御好評を頂いたが，あまりにも“遊び”や“ゆとり”がなくて多くの箇所で“行間を読む”ことを読者に求めているとのお小言も頂戴した．これには私もまったく同感であって，いつかは岩波版「一般線形代数」を土台にして，ページ数の制約にあまり囚われずに——そしてこれは読者の方には申し訳ないことであるが——価格にも囚われずに，少々の“遊び”と“ゆとり”も許してもらった

“大増補改訂版”を作製してみたいものだと夢見ていた．

ちょうどその頃，朝倉書店編集部の方にお会いする機会があり，私の夢をお話したところ関心を示して下さったので，岩波書店で上記の出版のお世話になっていた吉田宇一様とも御相談した上で，この増補改訂版を「線形代数汎論」というやや古風な表題をつけて朝倉書店から刊行していただくことにした．岩波版「一般線形代数」を読み直して気の付く所に赤字を入れて行ったら瞬く間に真っ赤になったのは言うまでもない．第一章として線形代数の周辺の話題（その誕生に貢献した人達のことも私の専門ではないが出来るだけ調べて含めた）を追加したし，幾つかの節，項の追加，組み直しもした．補章も追加して，そこでは，組合せ論，グラフ理論に関連した話題を積極的に論じた．その目的は，行列式の関連でこのような観点が本当は大切なのに線形代数の教科書などでは何となく補助的なもののように扱われていた順列や置換の格上げをすること，そして最近再び注目されるようになった“行列式や Pfaffian の除算を用いない計算法”の紹介もその一環として行うことにある．

これに関連して想起されるのは，その原稿が書かれてからもう六十数年が経つといわれる Феликс Рувимович Гантмахер（Feliks Ruvimovich Gantmakher, 1908-1964）の大著 Теория Матриц のことである．この本は，およそ行列の理論に関連する研究をする者にとっては必読の定本であるとされ，欧米各国語にも翻訳され（英訳 F. R. Gantmacher, The Theory of Matrices, I, II（Chelsea 1959）； 独訳 Matrizentheorie； 等々），ロシアでは版を重ねて 2004 年には第 5 版が Физматлит から刊行されている．応用分野も，力学，微分方程式，制御理論，数値計算，など，広くカバーしている．本書を著すに当たって，当然私の頭にはこの本のことがこびりついて離れなかったということを告白しなければならない．しかし，半世紀以上経つと，科学技術の進歩もあり，行列の理論自身の発展もさることながら，その応用の形態はすっかり様変わりしている．伝統の良さは保ちながら Гантмахер の心を現代に移したらどうなるかと自問しながら，仕事をしていた．最後の章の行列束のところがやや尻切れトンボになってしまったのは，私の力不足のためである．システムの安定論に深く関係する Routh-Hurwitz の理論（その延長線上の Харитонов の定理なども）を省いたのは，私の前著「数値計算―方程式の解法」（朝倉書店，1981）に詳しいからである．Pfaffian

については，線形代数の通り一遍の本にはあまり書かれていないが，最近また関心が高まっている主題でもあり，なるべく伝統的古典的な行列・行列式の理論と馴染むような形で扱うべく工夫した積りである．Гантмахер には含まれていない話題もかなり多く取り入れざるをえなかったのは時代の趨勢と感じた．

いろいろな事実を述べるに当たって理論的な証明と構成的な方法とがある場合，なるべく両方とも書くように努めたが，どちらか一方にせざるを得ないときには後者を優先した．特に，あまり一般の読者には馴染みがないかもしれない“交代化演算”をふんだんに利用（乱用？）したのもそのような意図からである．反変，共変のような“テンソル”的な概念は，線形代数においても本当は大切なもので，双対の概念の表記などもすっきりするのであるが，そうすることによって生じる読者側からの拒絶反応をいささか恐れてあまり表に出さなかった．もちろん，関連した事項はテンソル的な視点から見ても正しく書かれていると思う．“演習問題”には，普通の演習問題もあるが，単なる演習問題を超えて内容的には本文の一部に含めたほうがよいようなものも少なくない．そのような意味で，本書の演習問題は多様である．

この“まえがき”の冒頭にも記したように，数学の利用者のための現代的な線形代数の，私から見るとまだやや控え目な，本書が，数理的な仕事に携わる多くの人々のお役に立てれば幸いである．

あまりにも多数になるので，逐一お名前を挙げるのは差し控えるが，私の研究者生活を通して，ご指導を賜った先生方，諸先輩，共同研究を楽しむことができた同僚，後輩の方々，また私の研究や著作にいろいろ率直な御意見，御批判を賜った多くの方々に，この場を借りて御礼申し上げます．ただ，岩波講座の校正刷り全部に眼を通して貴重な多くの注意をしてくださった今井敏行君（現和歌山大学准教授）および朝倉版の校正刷りを綿密に点検して多くの重要な御助言を賜った東京大学の室田一雄教授とその研究室の方々にはお名前を挙げて感謝の意を表しなければならない．これも言うまでもないことかもしれないが，朝倉書店編集部には校正の過程で形式的な整合性，それを通じて内容にまでわたる改善の御示唆を数多く頂いた．記して御礼を申し上げたい．

顧みれば，岩波講座「応用数学」に関わっていたのは私の本務がもっとも忙しいときであった．二つの大学（東京大学工学部計数工学科，中央大学理工学部情

報工学科）を定年で退職した後はやや時間的には余裕ができた（はずである？）ので，本書の準備作業は主として自宅で行ったが，一部は非常勤で勤めていた（財）日本測量調査技術協会の会長室も使わせていただいた．後者の本務はこのような原稿書きではなかったが，大目に見逃してくださった同協会には感謝の他ない．

2009年7月吉日

伊理正夫

目　　次

1
線形代数の周辺

現在“線形代数(学)”と呼ばれている分野に関連した非技術的なことがらを，次章以降の技術的な話に入る前にいくつか述べておこう．この種のことは，“あとがき”や“囲み記事”の形にした方がよいのかもしれないが，本書では予備知識として知っていて損のない事項を，あえて冒頭におくことにした．したがって，本章は読みとばしていただいて差し支えない．

1.1　行列，行列式，線形代数小史

現在日本の大学で“線形代数”という名で教えられていることの内容は，半世紀ほど前には“代数学と幾何学”という科目名になっていたし，もう少し前には“行列(と行列式)の理論”と呼ばれていたこともあった．しかし，呼び名はともかく，本質的にはそれほど変化はないようである．もちろん，大学における講義も，教科書も，参考書も，それぞれの担当者や著者の創意工夫が凝らされていて，外見はずいぶん異なっていることもあるが．これは諸外国においてもほぼ同じような状況である．

行列，行列式，線形代数などの概念やそれを扱う数学的技術は，数学の他の分野にも必須なものであるばかりでなく，統計学，物理学，経済学，工学などの諸分野において重要な役割を果たすことが，ここ百数十年の間に次第に明らかになってきた．その歴史を詳述するのが本書の目的ではないので，ごく大ざっぱに流れを辿ってみよう．

最初に現れたのは“行列式”の概念であった．これは自然にみえる．いまの言葉で言えば，連立一次方程式あるいは連立代数方程式の中から変数を次々と消去して行くと，最後に残る1変数の式の係数は，もとの方程式の係数の多項式ある

いは有理式になる．それが行列式あるいは終結式と呼ばれるものである．(本書第4章の注意4.2.1，例4.3.1を見よ.)日本では関孝和(セキ・タカカズ，1642～1708)が1683年頃に終結式の考え方に，ヨーロッパではGottfried Wilhelm (von) Leibniz(ゴットフリート・ヴィルヘルム・ライプニッツ，1646～1716)がほぼ同時の1682年頃に行列式の概念に到達したといわれている．

“行列”あるいは行列代数と呼ばれる概念を表面に出して論じ始めたのは，Arthur Cayley(アーサー・ケイリー，1821～1895)とJames Joseph Sylvester(ジェームズ・ジョーゼフ・シルヴェスター，1814～1897)[1]の両人であるとされている．両人はほとんど同年代人であっただけでなく，ケンブリッジ大学の優等卒業試験(Tripos)に高成績で合格し(Sylvesterは1837年に2位で[2]，Cayleyは1842年に1位で)，若い頃は数学者としての職はなく，生計を立てるために法廷弁護士や記録係などとして働いていた．この間に両人は数学のあらゆる分野について語り合う機会をもったが，その中に行列の代数，群論，不変式論，幾何学などの諸問題が含まれていた．Cayleyは1863年にケンブリッジ大学の純粋数学の教授に任命された．給料は法律家のときより激減したが，数学を本職とできることに満足していた．彼はその後の人生をほとんどケンブリッジ大学で過ごし，1883年に英国の科学振興協会の会長にもなった．1882年の1月から5月までアメリカのジョンズ・ホプキンズ大学に滞在して講義をしたが，そこではSylvesterが数学の教授をしていた．Sylvesterは，Cayleyとは正反対の，幼い頃から激しい気性の人間でそのためにいろいろな問題を起こしたが，1854年にウリッジの英国陸軍士官学校の数学の教授になり，そこを1869年に55歳で定年退職した．そこでの教育の負担が大きいので数学者としての才能が磨耗されると不満をもち，学校には馴染めなかったらしい．1866年にはDe Morgan(ド・モルガン)の後を継いでロンドン数学会の第2代会長になった．1877年にはアメリカのジョンズ・ホプキンズ大学の教授を引き受け，1878年にアメリカで初の数学誌*American Journal of Mathematics*を立ち上げた．この職は数学の大学教授として典型的な環境を与えてくれたが彼には満足でなかったらしい．1883年，

---

1) Sylvesterの両親の代までの家族の姓はJosephだったという．アメリカでの居住権を得るに当たって，ユダヤ人臭くない名前にするために兄がJosephをミドルネームにして姓をSylvesterとしたのにならって，彼も若いうちにJosephをミドルネームにしてSylvesterと名乗ることにしたといわれている．

2) しかしユダヤ教信者であったため，大学を“卒業”することはできなかった．

Sylvester 既に 68 歳のとき，彼はオックスフォード大学の幾何学の教授に任命されたが，10 年ほどでロンドンに戻ってしまった．Sylvester は数学の研究以外に詩の研究にも力を注ぎ，“詩の諸法則” という著作も自慢にしていたそうである．

Cayley や Sylvester とはほとんど独立になされた，Hermann Günter Grassmann（ヘルマン・ギュンター・グラスマン，1809～1877）の貢献も無視してはならない．Grassmann は現代の言葉でいう “線形ベクトル空間” やベクトルの外積のような考え方に到達し，“外積代数” の創始者ともいわれている．後に Elie Cartan（エリー・カルタン，1869～1951）の “外微分形式” が有名になるより半世紀以上も前に *Die Lineare Ausdehnungslehre*（1844）を完成し，それを力学，電磁気学，幾何学などに応用した．彼の業績が正当に評価されるまでには約百年の月日が必要だった．Grassmann はシュテッティン（当時はプロシャ，ドイツの町 Stettin；現在はポーランド領の Szczecin，シュチェツィン）のギムナジウム（中学・高校レベルの高等教育機関）に主として関係していて，ベルリン大学にも何回か出入りしたらしいが，それもギムナジウムの教師としての資格を得る，高めるためであった．数学以外に，彼は印欧語の古形の研究もしていたという．

Grassmann の外積代数は，線形代数の中でも重要な役割を果たすが，外微分形式を通じて微分幾何学や物理学の諸分野にも有用な道具となる．本書では，そのことを陽には取り上げないで，実質的には等価な役割を果たす “交代化演算” の形をふんだんに使うことにする．これは Jan Arnoldus Schouten（ヤン・アルノルデュス・スホウテン（スカウテン），1883～1971）や Dirk Jan Struik（ディルク・ヤン・ストロイク，1894～2000）のオランダ学派の流儀に従ったものである．（Struik は結局アメリカ滞在が長かった．）

行列や行列式に関する研究成果が急速に累積してくると，それらをまとめた図書も著されるようになった．その走りは，行列式に関する Thomas Muir, *Theory of Determinants in the Historical Order of Development*（1883）であった．それまでに知られていたすべての結果をまとめたと言っており，当時はそれが可能だったということでもある．William H. Metzler によるその増補・改訂版 *A Treatise on the Theory of Determinants* が 1933 年に Longmans, Green and Co. から出版されたときには，766 ページの大著になっていた（Dover 版あり）．1933 年に Springer から出版された C. C. MacDuffee, *The Theory of Matrices* にはも

うかなり今風の行列代数の理論が展開されている(Chelsea 版あり).

現在では，行列や線形代数の既知の事実をすべて含んだ出版物など可能だとは思えない．ある立場から体系的に整理した参考書——本書もそれを目指したものであるが——，あるいは，初等的な入門書，教科書の類は多数あるが．線形代数やその応用に関する研究は，科学・技術の諸分野で広く展開されていて，研究や成果の動向が一望できるものではないが，“線形代数”に限った国際学術誌 *Linear Algebra and Its Applications* は 1968 年の創刊以来隆盛を極めている．また，“The International Linear Algebra Society”という国際学会(1989 年創設；1987 年に前身となる International Matrix Group が作られた)も活躍している．

なお，本節に登場した人物の履歴などについては(後章に登場する多くの人物も含めて)，かなり信頼性の高い http://www.gap-system.org/~history/Mathematicians/に拠るところが多い．また，Cayley と Sylvester については最近次の大著が出版されている．Tony CRILLY, *ARTHUR CAYLEY : Mathematician Laureate of the Victorian Age*, Johns Hopkins University Press, Baltimore, 2006, 784 pp. Karen Hunger PARSHALL, *JAMES JOSEPH SYLVESTER : Jewish Mathematician in a Victorian World*, Johns Hopkins University Press, Baltimore, 2006, 544 pp.

1.2　用語について

行列に関係した各種の用語について概観してみよう．

まず，“行列”という日本語は，“一列に並んでいるもの”のことを意味するのが普通である．科学技術用語にも“待ち行列(queue)”があるし，待ち行列理論(queu(e)ing theory)はオペレーションズ・リサーチや通信理論の中で重要な一部門をなしている．線形代数や行列式に関連して使われる“行列(matrix)”は，物を四角に(縦横に)並べたもので，縦の並びを“列(column)”，横の並びを“行(row)”と呼ぶところから，その“行”と“列”を併わせて“行列”と呼ぶことにしたものと思われる．因みに，英語などのヨーロッパ語では，matrix は“中に何かを入れる所(子宮など)”という意味であるから，物を四角に並べる枠のようなものと考えておいてよいのであろう．

行列，行，列などが，英，独，仏，露などの各国語で何と呼ばれているかを一覧しておく．

日：	行列	行	列
英：	matrix (*pl* matrices)	row	column
独：	Matrix (*f*; *pl* Matrizen)	Zeile (*f*; *pl* —n)	Spalte (*f*; *pl* —n)
仏：	matrice (*f*)	ligne (*f*)	colonne (*f*)
露：	ма́трица (*f*)	строка́ (*f*)	столбе́ц (*m*; —бца́, …)

お隣りの中国の数学用語は，同じ漢字圏ということで日本の用語と相互に影響しあっているが，行列関係では，一般の行列を“矩陣”，特に正方行列で“正方”を強調したいときには“方陣”というのはわかりやすい．単位行列が“単位方陣”だというので安心していると，逆行列は“反方陣”，直交行列は“正交矩陣”(これは“正交方陣”でもよさそうであるが)となる．固有値などは“特徴値”などとなっている．もっともこれは characteristic value という英語の訳だと思えば，日本語にその用例があっても不自然でないが．ところで，彼地には仮名がないので，ベクトルのことは“向量”と呼んでいるらしい．(テンソルのことは“張量”.)したがって，固有ベクトルは“特徴向量”となっている．中国本土と台湾での用語の差は，用いている漢字の書体の差だけであるようである．(厳密にいうと，日本語の漢字の書体は中国の繁体字とも簡体字ともいまや異なってしまっているが.)

線形代数，線形独立・従属というとき，“線形”と書くか“線型”と書くかについては，日本の中での意見の統一はまだ得られていないようである．本書では“線形”に統一してある．“形”は“かたち”で“型”は“かた”なので，微妙に意味が異なるから使い分けるべきであると拘る人も少なくない．中国では，これを“綫性”(綫は線の異体字)と呼んでいるので，“形”か“型”かの問題はすっきりしている．

現在では，行列の形にモデル化される物理や工学・経済の大型の問題が多い．それらの問題を“数値的に”解くために，いまや“線形計算(linear computation)”という分野が確立している．簡単に言えば，線形計算は“線形代数に関係した数値計算法”である．ところが，日本語で“線形計算”というと，正に，“(曲)線の形の計算法”という意味の，実用性の高い，そして歴史も長く需要も多い分野がある．たとえば，図 1.2.1 において，直線の道路を自動車で A まで

来て，そこから徐々にハンドルを切って点 B まで進み，そこでハンドルの角度を一定に保ち C まで進み，そこから徐々にハンドルを戻し，D でハンドルを戻し切って，そこからはまた直線に沿って進む，というようなとき，どのように道路の線の形を設計すればよいか．一定のハンドル角では自動車は一定の曲率の円弧を描いて走る．ハンドルを切らずに進むときには，曲率 0 の曲線，すなわち直線で進む．その間を“滑らかに”繋ぐにはどうしたらよいか，というのが線形計算の問題である．“なるべく急ハンドルを切らないように”ということを“ハンドルの角度を変えるときは角度の変化が進行距離に比例するように”と定式化する．曲線上の点の x-y 座標を (x, y)，ある点から曲線に沿って測った長さを s，点 (x, y) における曲線の接線の x 軸との角度を θ とすれば(図 1.2.2)，$\mathrm{d}x/\mathrm{d}s=\cos\theta$，$\mathrm{d}y/\mathrm{d}s=\sin\theta$，$\mathrm{d}\theta/\mathrm{d}s=$ 曲率 $=1/R=c\cdot$[ハンドルの角度] ($R=$ 曲率半径)であるから，曲線が直線である部分においては $\mathrm{d}\theta/\mathrm{d}s=0$ ($\theta=$const.)，円弧

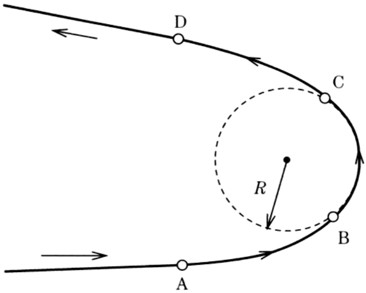


図 1.2.1　クロソイド曲線の例

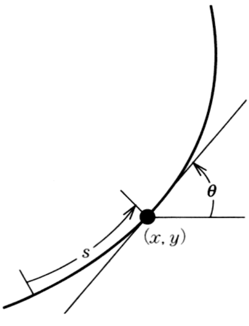


図 1.2.2

である部分においては $d\theta/ds=\text{const.}$，直線と円弧を繋ぐ部分においては，繋ぎ目のところから s を測ることにして，$d\theta/ds=\pi s/a^2$（a^2 は適当に選んだ定数），すなわち $\theta=(\pi/2a^2)s^2$ となる．そこで，この最後の部分の形は

$$x=\int_0^s \cos\frac{\pi s^2}{2a^2}\,ds, \qquad y=\int_0^s \sin\frac{\pi s^2}{2a^2}\,ds \tag{1.2.1}$$

となる．これは**クロソイド曲線**(clothoid curve)とか**緩和曲線**とか呼ばれる曲線であり，光の回折理論などに現れる**コルニュの螺線**(Cornu spiral)と同じものである．((1.2.1)の二つの積分は**フレネル積分**(Fresnel integral)と呼ばれているものである．）道路の線形の設計の分野では s を L と記し $a/\sqrt{\pi}$ を A と記して，クロソイドの形を

$$LR=A^2 \tag{1.2.2}$$

と表すことが多い．(A は長さの(物理)次元をもつ定数．）道路幅が広く多車線の道路の場合には上記の話はどうなるかとか，曲がりのために生じる遠心力を打ち消すのに道路幅方向につける傾斜(cant)をどう決めたらよいかとかいうような，“実際的な”問題はいろいろあるが，もちろんここではこれ以上立ち入らない．

行列代数の始祖の一人である A. Cayley(→1.1節)が，歿する直前に小冊子 *The Principle of Book-keeping by Double Entry*（複式簿記の原理）を著していることにも注目したい．彼は，「複式簿記はユークリッドの“比の理論”と同じように完全なものである」と述べているというが，このことは本書注意 2.2.2 とも関係するが，詳しくは論じない．いわゆる“行列簿記”の“行列”が線形代数で扱う行列と関係があるのかないのかについても．

1.3 人名について

どんな分野についても同じことであるが，本書でも多くの学者・研究者の名前が出てくる．ある人名のついた特殊な形の行列とか，定理とか，方法とかが，その人とどのような関りをもっているのかについては，いろいろと議論がある場合が少なくない．

たとえば 2.16.1 項の Vandermonde 行列(式)は実は他の人が既に用いていたとか，4.4 節に現れる Gauss, Jacobi, Seidel などの名は正当に使われているかと

か，その他議論をし出せばきりがない．これは線形代数関連のことに限らない．よく知られた例では，ベクトル解析などによく現れる Gauss の定理：

$$\iiint_V \operatorname{div} \boldsymbol{u}\, \mathrm{d}V = \iint_{\partial V} \boldsymbol{u} \cdot \mathrm{d}\boldsymbol{S} \tag{1.3.1}$$

がある．($\boldsymbol{u}$ はベクトル場，V は 3 次元空間のある領域，$\mathrm{d}V$ は体積要素，$\operatorname{div} \boldsymbol{u}$ は $\boldsymbol{u}$ の発散，∂V は V の境界面，$\mathrm{d}\boldsymbol{S}$ はその面積要素(ベクトル)である．詳しくはベクトル解析などの書物を参照のこと．）我国では，これを“Gauss の定理”と呼んでいる(そう呼ばないと通じない)が，ロシア(ソヴィエト連邦)では“Ostrogradskiĭ の定理”と呼び，英米・独で Gauss としか呼ばないことを非難さえしている——仏ではときどき Ostrogradskiĭ の名も入れることがあると言いながら．(Михаил Васильевич Остроградский：ミハイル・ヴァシリエーヴィッチ・オストログラツキー，1801～1862(暦法により 1861 とされることもある)．)

本書にもよく現れる“Cayley-Hamilton の定理”も，人によって“Hamilton-Cayley の定理”と呼ぶこともある．

このようなことに凝ると際限がないから，本書ではよほどのことがない限り，たとえそれが誤用であっても，我国で習慣的に行われている呼び名を用いることにする．

人名についてもう一つ注意しておかなければならないのは，日本語(片仮名表記)での表し方である．本書では，原則として原綴(英字が原綴でないものについては英字に転写)を用い，初出の所に片仮名で近似的な読み方を示すことにする．“読み”については，なるべく“原地主義”に従い“英語読み”は避けるように努める．しかし，昔も今も数学者の活躍は国際的で，ある人の名前の生誕地での読みと，主として活躍していた場所での読みが異なる場合が少なくない．また，人名の綴り自体も，同一人が異なる言語で書いた論文に異なる綴りを用いる例があるので話が複雑で，簡単に割り切れない．

1.4　記号について

本書で用いる記号については，その場その場で定義や説明をするが，転置行列と複素共役[1](Hermite 共役)行列，逆行列の表記についてだけあらかじめ述べておく(→ 2.3.1 項)．

行列 A の転置行列は，多くの書物，論文などでその表記はまちまちで，A'，${}^tA$，${}^{\mathrm{t}}A$，A^t，A^{t} とか書かれるが，本書では A^{T} と記すことにする．

同じく A の転置複素共役は，$A^*=\bar{A}'$ などと書かれる場合もあるが，本書では A^{H} と書くことにする．

逆行列 A^{-1} を A^{I} と記すこともあるが，本書では A^{-1} を採る．

1) ［キョウヤク］と読む．古い形が共軛であった．［キョウエキ］とは読まないように．

2
行列と行列式

本章ではごく古典的な接近法で，行列と行列式に関する基本事項を述べる．もちろん，次章のベクトル空間の概念を知ってから再び本章を見直すと，新しい視野が開けるのではあるが，いわゆる組合せ論的手法を用いる本章のようなやり方は，本格的な代数系(体や Euclid 環など)以外のものを対象とする場合にも通用する(かどうかが容易にわかる)ので，そのつもりで読んでいただきたい．

2.1 行　　　列

たとえば，実数体$\mathbb{R}$の上の2×3型の行列Aとは，6個の実数$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}$を次のように長方形に配列したものをいう：

$$A=\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}. \tag{2.1.1}$$

一般に，ある代数系Kの上の$m\times n$型の(“m行n列の”ともいう；m, nは自然数)**行列**(matrix)

$$A=[a_{ij}] \qquad (i=1,\cdots,m;\ j=1,\cdots,n)$$

も同様にして定義される：

$$A=\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}, \qquad a_{ij}\in K. \tag{2.1.2}$$

行列Aの横の並び$[a_{i1}\cdots a_{in}]$を**行**(第i行，あるいは行i)(row)，縦の並び

$$\begin{bmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

を**列**(第 j 列，あるいは列 j)(column)と呼ぶ．行列 A を構成する要素 a_{ij} のことを，A の (i,j) **要素**(element，entry，**成分**，component ともいう)と呼ぶ．

$n\times n$ 型の行列を，**次数**(order) n の**正方行列**(square matrix)という．

以降，K は有理数体 $\mathbb{Q}$，実数体 $\mathbb{R}$，複素数体 $\mathbb{C}$，整数環 $\mathbb{Z}$，あるいはそれらを係数とする 1 変数多項式環 $K[x]$(さらにはその商体である有理式体 $K(x)$)であるとして話を進める．それ以外の場合，たとえば K が分配束である場合などにも一部の話は適用できる．(代数系の諸概念については，たとえば，伊理・藤重：「応用代数」，コロナ社，1996，1988 など参照．)

すべての要素が 0 である $m\times n$ 型行列を $O_{m,n}$ と記し(m と n が前後関係から明らかなときは単に O と記す)，**零行列**(null matrix)と呼ぶ．

注意 2.1.1　より一般的，形式的には，行列 A は，二つの有限集合 I,J(空集合でもよい)と代数系 K に関して，$A:I\times J\to K$ なる関数のことであると定義することもできる．(I が行集合，J が列集合，$A:(i,j)\mapsto a_{ij}$ である．)実際，行や列の順序は応用上無意味で，単に便宜的に任意に定めたものにすぎないことが多い．さらに，I と J との間に対応関係があるとは限らない．

たとえば，I が原材料の種類の集合，J が製品の種類の集合，a_{ij} が製品 j を単位量作るのに要する材料 i の量であるというモデルのことを考えてみられたい．

I と J が同数の元からなるとき，すなわち $|I|=|J|$ のとき A は正方行列である．さらに，I と J の間に一対一対応があるとき，簡単にいえば $I=J$ であるとき，行列 A の (i,i) 要素のことを A の**対角要素**(diagonal element)，その他の要素を**非対角要素**(off-diagonal element)という $(i=1,\cdots,m(=n))$．$I=J$ の例としては，$I=J$ が n 個の都市の集合，a_{ij} が都市 i から j への距離(道路に沿っての距離，時間距離など)であるいわゆる「距離行列」がある．

I および J に(完全な)順序が定められている(本質的な意味のある順序のこともあるし，仮に定めたものであることもある)場合，$\{a_{ij}\mid i\geqq j\}$ を A の**下三角部分**(lower triangular part)，$\{a_{ij}\mid i\leqq j\}$ を**上三角部分**(upper triangular part)，下三角部分以外の要素がみな 0 である行列を**下三角行列**，上三角部分以外の要素がみな 0 である行列を**上三角行列**という．

また，$i>j+1$ のとき $a_{ij}=0$ である行列を上 **Hessenberg**(ヘッセンベルク)**行列**，$j>i+1$ のとき $a_{ij}=0$ であるものを下 Hessenberg 行列と呼んでいる．

非対角要素がみな 0 である(正方)行列を**対角行列**(diagonal matrix)といい，

$A=\mathrm{diag}[a_{11},\cdots,a_{nn}]$のように記す．$\mathrm{diag}[1,\cdots,1]$を$n$次の**単位行列**(unit matrix, identity matrix)と呼び，I_nと記す(明らかなときはnを略す)．

$|i-j|>b\ (\geqq 1)$のとき$a_{ij}=0$であるような行列$A=[a_{ij}]$は，**帯幅**(width)$2b+1$の**帯行列**(band matrix)と呼ばれる．特に，$b=1$のとき**三重対角**(tridiagonal)**行列**という．三重対角行列は，上 Hessenberg 行列であり，かつ下 Hessenberg 行列である．

> **注意 2.1.2**　特に，$m\times 1$型の行列のことを(m次元)**縦ベクトル**(**列ベクトル**)，$1\times n$型の行列のことを(n次元)**横ベクトル**(**行ベクトル**)と呼ぶ．また，1×1型の行列を**スカラー**(scalar)と呼び，それをKの要素a_{11}と区別しないこともあるが，本来は異なる概念である．

2.2　行　列　算

2.2.1　和　と　差

同じ型($m\times n$型)の二つの行列A,Bに対しては，それらの和$S=A+B$，差$D=A-B$が，対応する要素ごとの和，差として定義される．すなわち，$A=[a_{ij}],B=[b_{ij}],S=[s_{ij}],D=[d_{ij}]$とすれば，

$$s_{ij}=a_{ij}+b_{ij},\qquad d_{ij}=a_{ij}-b_{ij}\qquad (i=1,\cdots,m\,;\ j=1,\cdots,n)\tag{2.2.1}$$

となる．SもDもA,Bと同じ型である．

2.2.2　積

$l\times m$型の行列$A=[a_{ij}]$と$m\times n$型の行列$B=[b_{ij}]$に対して，積$P=[p_{ij}]=AB$($A\cdot B$とも記す)は$l\times n$型行列で

$$p_{ij}=\sum_{k=1}^{m}a_{ik}\cdot b_{kj}\qquad (i=1,\cdots,l\,;\ j=1,\cdots,n)\tag{2.2.2}$$

により定義される．(Aの列の本数とBの行の本数が等しいことが大切！)

2.2.3　スカラー倍

$m\times n$型の行列$A=[a_{ij}]$とKの元λ(スカラーと呼ぶこともある；注意 2.1.2 も見よ)に対して，Aのλ倍$B=[b_{ij}]=\lambda A$は$m\times n$型の行列で

$$b_{ij}=\lambda a_{ij}\qquad (i=1,\cdots,m\,;\ j=1,\cdots,n)\tag{2.2.3}$$

により定義される．

同じ型の行列 A, B と $\lambda, \mu(\in K)$ に対して，$\lambda A+\mu B$ が(2.2.1)，(2.2.3)を何回か用いて定義できる．

2.2.4 行列算の諸公式

上記の演算が以下の性質を満たすことは容易に確かめられる．(K における加減算，乗算がもつ性質によっているので，K を体や環でない代数系に取るときにはその点に注意する必要がある；演習問題 2.1，2.2 参照)

(i) A, B, C が同じ型の行列，O がそれらと同じ型の零行列，λ, μ, 1, $0\in K$ として，

$$1A=A,\quad 0A=O,\quad A-A=O,\quad A-B=A+(-1)B, \tag{2.2.4}$$

$$(\lambda\mu)A=\lambda(\mu A)=\mu(\lambda A), \tag{2.2.5}$$

$$(A+B)+C=A+(B+C), \tag{2.2.6}$$

$$A+B=B+A, \tag{2.2.7}$$

$$(\lambda+\mu)A=\lambda A+\mu A, \tag{2.2.8}$$

$$\lambda(A+B)=\lambda A+\lambda B. \tag{2.2.9}$$

(ii) A, B, C, O(零行列)，I(単位行列)はそれぞれ下式の中の諸演算が定義されるような型の行列(型を右下付添字で明示する)であるとして，

$$O_{l,m}A_{m,n}=O_{l,n},\quad A_{l,m}O_{m,n}=O_{l,n}, \tag{2.2.10}$$

$$\lambda(A_{l,m}B_{m,n})=(\lambda A_{l,m})_{l,m}B_{m,n}=A_{l,m}(\lambda B_{m,n})_{m,n}, \tag{2.2.11}$$

$$A_{k,l}(B_{l,m}C_{m,n})=(A_{k,l}B_{l,m})C_{m,n}, \tag{2.2.12}$$

$$A_{l,m}(B_{m,n}+C_{m,n})=A_{l,m}B_{m,n}+A_{l,m}C_{m,n}, \tag{2.2.13}$$

$$(A_{l,m}+B_{l,m})C_{m,n}=A_{l,m}C_{m,n}+B_{l,m}C_{m,n}, \tag{2.2.14}$$

$$I_mA_{m,n}=A_{m,n}I_n=A_{m,n}. \tag{2.2.15}$$

注意 2.2.1 要するに，行列の間の演算は——型が適合してさえいれば——ほぼ普通の数の間の演算から類推される性質を有している(たとえば，いくつもの行列の和を取るときの括弧は必要ないし，積についても同様である；和は可換でもある)ということである．行列 A に対角行列 $\mathrm{diag}[\lambda,\lambda,\cdots,\lambda]=\lambda I$ を(左あるいは右から)掛けることは，(2.2.15)，(2.2.11)により，A を λ 倍することと等価である．なお，ここで特に注意しなければならないのは，AB と BA の両方が定義さ

れるときでも一般には $AB=BA$ とはならない(積は可換でない)ことである．たとえば，A が $m\times n$ 型，B が $n\times m$ 型なら，AB は $m\times m$ 型，BA は $n\times n$ 型であるし，A と B が共に n 次正方行列であっても，

$$A=\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix},\quad B=\begin{bmatrix}1 & 0\\ 1 & 1\end{bmatrix},\quad AB=\begin{bmatrix}2 & 1\\ 1 & 1\end{bmatrix},\quad BA=\begin{bmatrix}1 & 1\\ 1 & 2\end{bmatrix}$$

である．また $A\neq O$, $B\neq O$ でも $AB=O$ となることがある．たとえば，

$$A=\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix},\quad B=\begin{bmatrix}2 & -1\\ -2 & 1\end{bmatrix},\quad AB=\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix},\quad BA=\begin{bmatrix}3 & 3\\ -3 & -3\end{bmatrix}$$

である．

注意 2.2.2　$m\times n$ 型の行列 A の要素が物理量を表していて，A と他の行列との積が作られるような場合(意味のあるモデルでは必ずそうである)には，A の要素の物理次元は

$$[a_{ij}\text{ の物理次元}]=[\text{行 } i \text{ に対応する物理次元}]/[\text{列 } j \text{ に対応する物理次元}] \tag{2.2.16}$$

という形に表されなければならない．実際，積 AB の $(i,1)$ 要素は，

$$\sum_{j=1}^{n} a_{ij}b_{j1} \tag{2.2.17}$$

であるが，(2.2.17)の和が物理的に無意味でないためには，各項 $a_{ij}b_{j1}$ が同じ物理次元を有していなければならない．このことから(2.2.16)が導かれる．(より詳しくは読者が自ら考察されたい．"次元解析"と呼ばれる技法との関係も深い．)

注意 2.2.3　正方行列 A に対しては，積 A, $A\cdot A$, $A\cdot A\cdot A$, …が常に定義される．これらから和やスカラー倍により作られる行列が積に関して可換であることも明らかである．そこで，1変数多項式 $f(x)=a_0+a_1x+\cdots+a_mx^m$ に対して，A の**行列多項式**(matrix polynomial)

$$f(A)=I+a_1A+a_2A^2+\cdots+a_mA^m \tag{2.2.18}$$

を定義することができる．ここで，

$$A^0=I,\quad A^p=A^{p-1}\cdot A=A\cdot A^{p-1}\qquad (p=1,2,\cdots) \tag{2.2.19}$$

とする．このとき，$f(x)=g(x)h(x)$ ならば，$f(A)=g(A)h(A)$ となることも明らかであろう．(逆は一般には成り立たない．)

注意 2.2.4　n 次正方行列の全体を $\mathcal{M}_n$ と記す．$\mathcal{M}_n$ の元の間には，(可換な)和と(非可換な)積が定義されていて，(2.2.13)，(2.2.14)の形の分配則が成立する．そこで，$\mathcal{M}_n$ は一つの環(行列環，多元環)をなすことになる．

2.3 対称行列，Hermite 行列，置換行列，基本行列

2.3.1 行列の転置と Hermite 共役

$m\times n$ 型の行列 $A=[a_{ij}]$ の**転置**(transpose) $B=[b_{ij}]=A^{\mathsf{T}}$ とは，

$$b_{ij}=a_{ji} \qquad (i=1,\cdots,n\,;\ j=1,\cdots,m) \tag{2.3.1}$$

で定められる $n\times m$ 型の行列のことである．

$K=\mathbb{C}$(複素数体)のとき，$m\times n$ 型の行列 A の**共役**(conjugate) $C=[c_{ij}]=\bar{A}$ とは，

$$c_{ij}=\bar{a}_{ij} \tag{2.3.2}$$

で定められる $m\times n$ 型行列のことである($\bar{a}$ は a の共役複素数)．共役転置行列 $(\bar{A})^{\mathsf{T}}(=\overline{(A^{\mathsf{T}})}$ であることにも注意！)を A^{H} と記し，A の **Hermite**(エルミート)**共役**とも呼ぶ．

これらの定義から，以下の諸性質が成り立つことは明らかであろう．($\lambda\in K$ とする．共役転置に関するものは $K=\mathbb{C}$ のとき．)

(i) $(\lambda A)^{\mathsf{T}}=\lambda A^{\mathsf{T}}$，$(\lambda A)^{\mathsf{H}}=\bar{\lambda}A^{\mathsf{H}}$. (2.3.3)

(ii) $(A+B)^{\mathsf{T}}=A^{\mathsf{T}}+B^{\mathsf{T}}$，$(A+B)^{\mathsf{H}}=A^{\mathsf{H}}+B^{\mathsf{H}}$. (2.3.4)

(iii) $(AB)^{\mathsf{T}}=B^{\mathsf{T}}A^{\mathsf{T}}$，$(AB)^{\mathsf{H}}=B^{\mathsf{H}}A^{\mathsf{H}}$. (2.3.5)

(iv) $(A^{\mathsf{T}})^{\mathsf{T}}=A$，$(A^{\mathsf{H}})^{\mathsf{H}}=A$. (2.3.6)

正方行列 A は，

$A^{\mathsf{T}}=A$ のとき，**対称**(symmetric)行列，

$A^{\mathsf{T}}=-A$ のとき，**反対称**(anti-symmetric)行列，**交代**(alternating)行列あるいは**歪対称**(skew-symmetric)行列，

$A^{\mathsf{H}}=A$ のとき，**Hermite** 行列，

$A^{\mathsf{H}}=-A$ のとき，**歪 Hermite**(skew-Hermite)行列

という．(形容詞として用いる Hermite は Hermitian，hermitian と書くこともある．)

注意 2.3.1 上記の定義は，行集合 I と列集合 J との間に一対一の対応関係がある($I=J$ とみなせる)場合しか意味をもたない(→注意 2.1.1 を参照)．一方，$I=J$ であるような任意の(正方)行列 A に対して，$B=\dfrac{1}{2}(A+A^{\mathsf{T}})$ は対称行列，また，

$C=\frac{1}{2}(A-A^{\mathrm{T}})$は反対称行列である(式(2.2.7)，(2.3.3)，(2.3.4)，(2.3.6)による)．そこで，Aはその対称部分と反対称部分に分解される：

$$A=\frac{1}{2}(A+A^{\mathrm{T}})+\frac{1}{2}(A-A^{\mathrm{T}}) \tag{2.3.7}$$

この分解が一意的であることは明らかであろう．

ほとんど同様にして($K=\mathbb{C}$のとき)，

$$A=\frac{1}{2}(A+A^{\mathrm{H}})+\frac{1}{2}(A-A^{\mathrm{H}}) \tag{2.3.8}$$

という，Hermite 部分と歪 Hermite 部分への分解もできる．

定義から直ちにわかるように，反対称行列の対角要素は 0，Hermite 行列の対角要素は実数，歪 Hermite 行列の対角要素は純虚数である．Kの標数が 2 のときには，反対称行列は対角要素が 0 の対称行列であるということにしておく．

注意 2.3.2　Aが Hermite 行列なら iAは歪 Hermite 行列，Aが歪 Hermite 行列なら iAは Hermite 行列である．

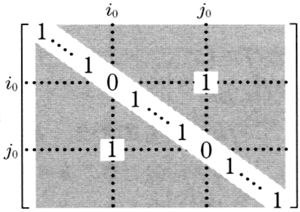
2.3.2　置換行列と基本行列

正方行列Pの要素がどれも 0 あるいは 1 に等しく，しかもどの行にも 1 がちょうど 1 個，どの列にも 1 がちょうど 1 個あるときに，Pは**置換**(permutation)行列と呼ばれる．m次の置換行列$P=[p_{ij}]$の行iにおける 1 が列j_iにあるとする($p_{ij_i}=1\,(i=1,\cdots,m)$, $p_{ij}=0\,(j\neq j_i)$)．このとき，$m\times n$型の行列Aに対して，$A'=[a'_{ij}]=PA$の行集合の順序はAの行集合の順序に置換[1)]

$$\begin{pmatrix} j_1 & j_2 & \cdots & j_m \\ 1 & 2 & \cdots & m \end{pmatrix}$$

を施すことによって得られる．また，$n\times m$型の行列Bに対して，BPはBの列集合の順序を置換したものになっている．特に，あるi_0, $j_0\,(i_0\neq j_0)$について，$p_{ii}=1\,(i\neq i_0,\ j_0)$, $p_{i_0j_0}=p_{j_0i_0}=1$, $p_{i_0i_0}=p_{j_0j_0}=0$, その他のi,jに対して$p_{ij}=0$であるような置換行列

1)　置換，偶置換，奇置換などについては補章 A.3 節も参照のこと．

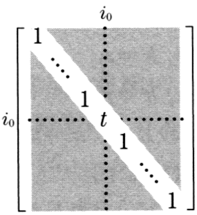
$P(i_0, j_0) \equiv$ 

（カゲをつけた部分は 0）

(2.3.9)

は**互換**(transposition)行列で，$P(i_0,j_0)$を行列 A の左(右)から掛けると，A の行 i_0 と行 j_0(列 i_0 と列 j_0)の入れ換えをひき起こす．

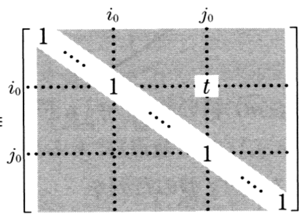
m 次対角行列 $T(i_0;t)=\mathrm{diag}[\overset{1}{1},\cdots,1,\overset{i_0}{t},1,\cdots,\overset{m}{1}]$　　$(t\in K)$：

$T(i_0;t) \equiv i_0$ 
（カゲをつけた部分は 0）

(2.3.10)

を $m\times n$ 型の行列 A に左から掛けたもの $T(i_0;t)A$ は，A の行 i_0(のすべての要素)を t 倍した行列に等しい．また，$BT(i_0;t)$は，B の列 i_0 を t 倍したものになる．

$E(i_0,j_0)$を，$e_{i_0j_0}=1$ で，それ以外の要素はすべて 0 の m 次正方行列とし，$W(i_0,j_0;t)=I_m+tE(i_0,j_0)$ $(t\in K,\ i_0\neq j_0)$とおく：

$W(i_0,j_0;t) \equiv$ 
（カゲをつけた部分は 0）.

(2.3.11)

$W(i_0,j_0;t)A$ は，A の行 j_0 を t 倍して行 i_0 に加えたものに等しい．また，$BW(i_0,j_0;t)$ は，B の列 i_0 を t 倍して列 j_0 に加えたものに等しい．

$P(i_0,j_0)$ $(i_0\neq j_0)$，$T(i_0;t)$ $(t\neq 0)$，$W(i_0,j_0;t)$ $(i_0\neq j_0)$の形の行列のことを**基本**(elementary)行列と呼ぶことがある(→2.15 節を参照)．直接容易に確認できるように，次のような関係がある($i_0\neq j_0$ とする)．

$$P(i_0,j_0)\,P(i_0,j_0)=I, \qquad T(i_0;t)\,T(i_0;1/t)=I, \qquad W(i_0,j_0;t)\,W(i_0,j_0;-t)=I. \tag{2.3.12}$$

$$P(i_0,j_0)\,T(k;t)=T(h;t)\,P(i_0,j_0) \quad \begin{pmatrix} k\neq i_0,j_0 \text{ のとき } & h=k \\ k=i_0 \text{ のとき } & h=j_0 \\ k=j_0 \text{ のとき } & h=i_0 \end{pmatrix}. \tag{2.3.13}$$

$$P(i_0,j_0)\,W(k,h;t)=W(k',h';t)\,P(i_0,j_0)$$
$$\begin{pmatrix} k,h\neq i_0,j_0 \text{ のとき } & k'=k,\ h'=h \\ k=i_0,\ h\neq j_0 \text{ のとき } & k'=j_0,\ h'=h \\ k=j_0,\ h\neq i_0 \text{ のとき } & k'=i_0,\ h'=h \\ k\neq i_0,\ h=j_0 \text{ のとき } & k'=k,\ h'=i_0 \\ k\neq j_0,\ h=i_0 \text{ のとき } & k'=k,\ h'=j_0 \\ k=i_0,\ h=j_0\ (k=j_0,\ h=i_0) \text{ のとき } & k'=h,\ h'=k \end{pmatrix}. \tag{2.3.14}$$

$$W(i_0,j_0;t)\,T(k;s)=T(k;s)\,W(i_0,j_0;t') \quad \begin{pmatrix} k\neq i_0,j_0 \text{ のとき } & t'=t \\ k=j_0 \text{ のとき } & t'=ts \\ k=i_0 \text{ のとき } & t'=t/s \end{pmatrix}. \tag{2.3.15}$$

$$P(i_0,j_0)=T(j_0;-1)\,W(i_0,j_0;1)\,W(j_0,i_0;-1)\,W(i_0,j_0;1). \tag{2.3.16}$$

2.4 部分行列，ブロック行列，Kronecker 積

2.4.1 部分行列と行列の分割

$m\times n$ 型の行列 $A=[a_{ij}]$ の $r(\leqq m)$ 本の行 $I'=\{i_1,\cdots,i_r\}$ $(i_1<\cdots<i_r)$ と $s(\leqq n)$ 本の列 $J'=\{j_1,\cdots,j_s\}$ $(j_1<\cdots<j_s)$（の交差点）にある要素を取り出して自然な順に配列することによってできる $r\times s$ 型の行列を

$$A_{J'}^{I'}=A_{j_1\cdots j_s}^{i_1\cdots\cdots i_r}$$

と記し，A の（行 $i_1,\cdots,i_r$，列 $j_1,\cdots,j_s$ からなる）**部分行列**(submatrix) と呼ぶ．（注意 2.1.1 の用語に従えば，部分行列は関数 $A: I\times J\to K$ の定義域を $I'\times J'$ に制限したものといえる．）

$m\times n$ 型の行列 A の行集合 I を r 個の部分 $I_1,\cdots,I_r$ に分け（$|I_k|=m_k$ とする），列集合 J を s 個の部分 $J_1,\cdots,J_s$ に分け（$|J_h|=n_h$），$A_{kh}=A_{J_h}^{I_k}$ と記すと，A

は自然に，

$$A=\begin{array}{c}\\ \updownarrow m_1\\ \updownarrow m_2\\ \vdots\\ \updownarrow m_r\end{array}\!\!\begin{array}{c}\leftarrow n_1\rightarrow \quad \leftarrow n_2\rightarrow \quad \cdots \quad \leftarrow n_s\rightarrow\\ \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s}\\ A_{21} & A_{22} & \cdots & A_{2s}\\ \vdots & \vdots & & \vdots\\ A_{r1} & A_{r2} & \cdots & A_{rs}\end{bmatrix}\end{array} \tag{2.4.1}$$

のように，いわば“部分行列を要素とする行列”のように表される．(2.4.1)のような表現を A の**分割**(partition)，$(m_1,\cdots,m_r;n_1,\cdots,n_s)$ を**分割の型**(partition type)という．このときには，部分行列 A_{kh} を**ブロック**(block)，(2.4.1)のような表現をブロックからなる一つの行列とみなして**ブロック行列**(block matrix)ともいう．

2.2 節で定義した行列の間の演算が，分割の型が適合しているならば，ブロック行列の間にも自然に拡張定義されることは，ほとんど自明であろう．たとえば，$l\times m$ 型の行列 A が $(l_1,\cdots,l_r;m_1,\cdots,m_s)$ という型に分割され，$m\times n$ 型の行列 B が $(m_1,\cdots,m_s;n_1,\cdots,n_t)$ という型に分割されているとき，$l\times n$ 型の行列 $C=AB$ は $(l_1,\cdots,l_r;n_1,\cdots,n_t)$ という型に分割され，それぞれの行列のブロックを $A_{ik},B_{kj},C_{ij}\,(i=1,\cdots,r;j=1,\cdots,t;k=1,\cdots,s)$ とすると

$$C_{ij}=\sum_{k=1}^{s}A_{ik}B_{kj}\qquad (i=1,\cdots,r;j=1,\cdots,t) \tag{2.4.2}$$

と表される．

ブロック行列の転置は，各ブロックごとに転置を作り，そしてブロックの配置を転置したものである．たとえば

$$A=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix}\text{に対して}\quad A^{\top}=\begin{bmatrix}A_{11}^{\top} & A_{21}^{\top}\\ A_{12}^{\top} & A_{22}^{\top}\end{bmatrix}$$

となる．

$m\times n$ 型の行列 A を，$(1,\cdots,1;n)$ という型に分割して，各行が作る横ベクトル(行ベクトル) $\boldsymbol{a}_{1\cdot},\cdots,\boldsymbol{a}_{m\cdot}$ の集まり

$$A=\begin{bmatrix}\boldsymbol{a}_{1\cdot}\\ \vdots\\ \boldsymbol{a}_{m\cdot}\end{bmatrix} \tag{2.4.3}$$

とみなしたり，$(m;1,\cdots,1)$ という型に分割して，各列が作る縦ベクトル（列べ

クトル) $\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot n}$ の集まり

$$A=[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot n}] \tag{2.4.4}$$

とみなしたりして扱うと便利なことが多い.

正方行列(で行集合 I=列集合 J のもの)の分割の型 $(n_1,\cdots,n_r;n_1,\cdots,n_r)$ を**対称な分割**(symmetric partition)と呼ぶ. 行列 A が対称な分割をされていて, そのブロックが, $k>h$ ($k<h$)のとき $A_{kh}=O_{n_k n_h}$ ならば, A を**ブロック上(下)三角**(block upper(lower) triangular)行列と呼ぶ. また, $k\neq h$ のとき $A_{kh}=O_{n_k n_h}$ ならば, A を**ブロック対角**(block diagonal)行列と呼び $\mathrm{diag}[A_{11},\cdots,A_{rr}]$ と記す. これらのブロック行列の**対角ブロック**(diagonal block) A_{kk} は, もちろん, n_k 次正方行列である.

2.4.2 Kronecker 積

$m_1\times n_1$ 型の行列 $A=[a_{ij}]$ と $m_2\times n_2$ 型の行列 $B=[b_{ij}]$ に対して, A と B の **Kronecker**(クロネッカー)**積**(Kronecker product) $A\otimes B$ とは, ブロック行列の形で表したとき,

$$A\otimes B=\begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n_1}B \\ a_{21}B & a_{22}B & \cdots & a_{2n_1}B \\ \vdots & \vdots & & \vdots \\ a_{m_11}B & a_{m_12}B & \cdots & a_{m_1n_1}B \end{bmatrix} \tag{2.4.5}$$

で定義される $(m_1m_2)\times(n_1n_2)$ 型の行列のことである. (注意 2.1.1 の用語によれば, $A: I_1\times J_1\to K$, $B: I_2\times J_2\to K$ に対して, $A\otimes B: (I_1\times I_2)\times(J_1\times J_2)\to K$ は, $A\otimes B: ((i_1,i_2),(j_1,j_2))\mapsto a_{i_1j_1}b_{i_2j_2}$ により定義されるものである. このことからもわかるように, $A\otimes B$ と $B\otimes A$ は, 行の順序と列の順序が異なるだけである.)

また, 定義から直ちに

$$\begin{aligned} &A^{\top}\otimes B^{\top}=(A\otimes B)^{\top}, \\ &(A_1+A_2)\otimes B=(A_1\otimes B)+(A_2\otimes B), \\ &A\otimes(B_1+B_2)=(A\otimes B_1)+(A\otimes B_2), \\ &(A_1\otimes B_1)\cdot(A_2\otimes B_2)=(A_1\cdot A_2)\otimes(B_1\cdot B_2), \\ &(A\otimes B)\otimes C=A\otimes(B\otimes C) \end{aligned} \tag{2.4.6}$$

などの関係式が得られる(行列の型は式が意味をもつようなものであるとする).

$m_1\times n_1$ 型の行列 $A=[a_{ij}]$ は，変数 $y_1,\cdots,y_{m_1}$ を変数 $x_1,\cdots,x_{n_1}$ の 1 次式

$$y_i=\sum_{j=1}^{n_1}a_{ij}x_j \qquad (i=1,\cdots,m_1) \tag{2.4.7}$$

で表すときの係数であると考えることができる．$m_2\times n_2$ 型の行列 $B=[b_{ij}]$ についても，変数 $v_1,\cdots,v_{m_2}$ を変数 $u_1,\cdots,u_{n_2}$ の 1 次式

$$v_i=\sum_{j=1}^{n_2}b_{ij}u_j \qquad (i=1,\cdots,m_2) \tag{2.4.8}$$

で表すときの係数であると考える．このとき，変数の積を新しい変数とみなして（これを**テンソル積**などと呼ぶこともあるが），$m_1\cdot m_2$ 個の変数

$$\begin{aligned}
&w_1=y_1v_1,\ w_2=y_1v_2,\ \cdots,\ w_{m_2}=y_1v_{m_2},\\
&w_{m_2+1}=y_2v_1,\ w_{m_2+2}=y_2v_2,\ \cdots,\ w_{2m_2}=y_2v_{m_2},\\
&\qquad\qquad\cdots\cdots\\
&\qquad\qquad\cdots\cdots\\
&w_{(m_1-1)m_2+1}=y_{m_1}v_1,\ w_{(m_1-1)m_2+2}=y_{m_1}v_2,\ \cdots,\ w_{m_1m_2}=y_{m_1}v_{m_2}
\end{aligned} \tag{2.4.9}$$

は，同じように作った $n_1\cdot n_2$ 個の変数

$$z_1=x_1u_1,\ z_2=x_1u_2,\ \cdots,\ z_{n_1n_2}=x_{n_1}u_{n_2} \tag{2.4.10}$$

の 1 次式で表されるが，その形が

$$w_i=\sum_{j=1}^{n_1\cdot n_2}(A\otimes B)_{ij}z_j \qquad (i=1,\cdots,m_1\cdot m_2) \tag{2.4.11}$$

となることは容易に知られる．Kronecker 積は，このように考えるとわかりやすい場合が多い．

注意 2.4.1 Kronecker 積は**直積**(direct product)と呼ばれることもある．
注意 2.4.2 固有値に関連したことは 5.5 節も参照のこと．

2.5 交代化演算

2.5.1 交代化の定義

$i_k, j_k\in\{1,2,\cdots,n\}$ とする．$\varepsilon^{i_1\cdots i_r}_{j_1\cdots j_r}$ を，

(i) $\{i_1,\cdots,i_r\}\neq\{j_1,\cdots,j_r\}$ であるか，$i_1,\cdots,i_r$ の中に同じものがあるか，あるいは $j_1,\cdots,j_r$ の中に同じものがあるとき，

$$\varepsilon^{i_1\cdots i_r}_{j_1\cdots j_r}=0 \tag{2.5.1}$$

であり，

(ii)　$i_1,\cdots,i_r$ がすべて異なり，$j_1,\cdots,j_r$ もすべて異なり，かつ $\{i_1,\cdots,i_r\}=\{j_1,\cdots,j_r\}$ であるとき，

$$\begin{pmatrix} i_1 & \cdots & i_r \\ j_1 & \cdots & j_r \end{pmatrix} \text{が偶置換なら}\quad \varepsilon^{i_1\cdots i_r}_{j_1\cdots j_r}=1, \tag{2.5.2}$$

$$\begin{pmatrix} i_1 & \cdots & i_r \\ j_1 & \cdots & j_r \end{pmatrix} \text{が奇置換なら}\quad \varepsilon^{i_1\cdots i_r}_{j_1\cdots j_r}=-1 \tag{2.5.3}$$

であると定義する．(補章 A.3 節も参照．)

$i_1\cdots i_r$ の順に添字が並んでいる式(これら以外に添字を有していてもよい) $F_{i_1\cdots i_r}$ に対して，$F_{i_1\cdots i_r}$ をこれらの添字に関して**交代化**する(to alternate)というのは，$F_{i_1\cdots i_r}$ から

$$F_{[i_1\cdots i_r]}\equiv\frac{1}{r\,!}\sum_{j_1=1}^{n}\cdots\sum_{j_r=1}^{n}\varepsilon^{j_1\cdots j_r}_{i_1\cdots i_r}F_{j_1\cdots j_r} \tag{2.5.4}$$

を作ることをいう．すなわち，$F_{[i_1\cdots i_r]}$ は，$i_1,\cdots,i_r$ がすべて異なるとき(それ以外の場合は $F_{[i_1\cdots i_r]}=0$ となる)，$(i_1\cdots i_r)$ の置換 $(j_1\cdots j_r)$ を添字にもつ式 $F_{j_1\cdots j_r}$ を，置換が偶置換なら正の符号，奇置換なら負の符号をつけて加え合わせたものを，置換の総数 $r\,!$ で割ったものである．たとえば，

$$F_{[122]}=0,$$

$$F_{[123]}=\frac{1}{6}(F_{123}+F_{231}+F_{312}-F_{213}-F_{132}-F_{321}),$$

$$F_{1[23]3}=\frac{1}{2}(F_{1233}-F_{1323})$$

となる．なお，式の中で連続していない添字に関して交代化をおこなうときには，交代化に関係しない添字を | | で囲む．たとえば，

$$F_{[12|34|5]}=\frac{1}{6}(F_{12345}+F_{25341}+F_{51342}-F_{21345}-F_{15342}-F_{52341}).$$

交代化の手続きの表記が混乱しないように，行列 $A=[a_{ij}]$ の (i,j) 要素を a^i_j あるいは a^j_i と記すこともある．

$G_{i_1\cdots i_r}=F_{[i_1\cdots i_r]}$ なら，$G_{i_1\cdots i_r}$ は添字 $i_1,\cdots,i_r$ に関して**交代的**(alternating)である，すなわち，これらの添字の任意の2個を入れ替えると G の符号が変わる．$F_{i_1\cdots i_r}$ がその添字に関して交代的なら，$F_{[i_1\cdots i_r]}=F_{i_1\cdots i_r}$ である．

また，明らかに，$r>n$ なら．$F_{[i_1\cdots i_r]}=0$ である．

2.5.2 交代化演算の諸性質

交代化の演算は以下の性質を有する．

(i) $a^{[i_1}_{j_1}\cdots a^{i_r]}_{j_r}=a^{i_1}_{[j_1}\cdots a^{i_r}_{j_r]}=a^{[i_1}_{[j_1}\cdots a^{i_r]}_{j_r]}$. (2.5.5)

(ii) $\sum_{i_1=1}^{n}\cdots\sum_{i_r=1}^{n}F_{[i_1\cdots i_r]i_1\cdots i_r}=\sum_{i_1=1}^{n}\cdots\sum_{i_r=1}^{n}F_{i_1\cdots i_r[i_1\cdots i_r]}=\sum_{i_1=1}^{n}\cdots\sum_{i_r=1}^{n}F_{[i_1\cdots i_r][i_1\cdots i_r]}$. (2.5.6)

(iii) $G_{i_1\cdots i_r i_{r+1}\cdots i_s}=F_{[i_1\cdots i_r]i_{r+1}\cdots i_s}$ のとき　$G_{[i_1\cdots i_s]}=F_{[i_1\cdots i_s]}$. (2.5.7)

(iv) $G_{i_1\cdots i_s}=F_{[i_1\cdots i_s]}$ のとき　$G_{[i_1\cdots i_r]i_{r+1}\cdots i_s}=G_{i_1\cdots i_s}$. (2.5.8)

(v) $\sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n}F_{[i_1\cdots i_n][i_1\cdots i_n]}=\sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n}F_{[i_1\cdots i_n]i_1\cdots i_n}$

$=\sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n}F_{i_1\cdots i_n[i_1\cdots i_n]}=n!\,F_{[1\cdots n][1\cdots n]}$. (2.5.9)

略証　以下の略証はあくまでも“略証”である．$r=3$ くらいに対して具体的に式を書き下してみれば素直に納得できよう．

(i)　$r!$ 倍して眺めると，最左辺は$(i_1\cdots i_r)$のすべての置換に符号をつけて加え合わせたものであるが，その和の各項ごとに a^i_j の積の順序を並べ換えて $i_1\cdots i_r$ が元のままの順になるようにすると，$j_1\cdots j_r$ の方が置換を受けたのと同じことになる．このとき置換の偶奇は等しいから，中辺の表式を得る．最左辺の式が添字 $j_1,\cdots,j_r$(中辺の式が添字 $i_1,\cdots,i_r$)に関して交代的であることも明らかであるから，最右辺が他の表式に等しいことがわかる．

(ii)　やはり $r!$ 倍して眺める．i_k は $1,\cdots,n$ の範囲を動いて和を取る添字であるから，他の添字と区別がつけられる限り，別の記号を用いてもよい．最左辺で前半の $i_1,\cdots,i_r$ の置換をしたら，その後で F の前半の添字が見掛け上元のままになるように記号の書き換えをすると，後半の添字が見掛け上置換を受けたようになる．$\sum$ を取る順序も変更する(してよい！)と，中辺の表式を得る．

(iii)　$r!F_{[i_1\cdots i_r]i_{r+1}\cdots i_s}$ を定義通りに書き下してから $r!G_{[i_1\cdots i_s]}$を作ると，書き下した式の各項から，$F_{[i_1\cdots i_s]}$に(符号も含めて)等しい式が出る．書き下した項の数は $r!$ であるから，結局 $F_{[i_1\cdots i_s]}=G_{[i_1\cdots i_s]}$となる．

(iv)　$G_{i_1\cdots i_s}$ は $i_1,\cdots,i_s$ に関して交代的であるから，もちろん，$i_1,\cdots,i_r$ に関しても交代的である．

(v)　最右辺以外は，(ii)の関係式の特別な場合である．この関係式の特徴は，交代化に係る添字の数が添字が動く範囲と等しいという点である．$\sum$ を取るとき，i_k の中に同じ値が現れるような組合せは$[i_1\cdots i_n]$を和に展開したときに互い

に消し合うから，和に最終的に残る項は最右辺を展開したときの形のもの($F_{k_1\cdots k_n h_1\cdots h_n}$，$(k_1\cdots k_n)$も$(h_1\cdots h_n)$も$(1\cdots n)$のある置換)しかない．各項の符号が$\varepsilon_{k_1\cdots k_n}^{1\cdots n}\varepsilon_{h_1\cdots h_n}^{1\cdots n}$に等しいこともわかる．さらに，同じ形の項は$1,\cdots,n$を$i_1,\cdots,i_n$に割り当てる割り当て方の数だけ重複して現れるから，最左辺，中辺の和の中には$F_{[1\cdots n][1\cdots n]}$が$n!$回重複して現れる．

注意 2.5.1　上記の(iii)，(iv)は，交代化を続けておこなうときに，その範囲の一方が他方を含んでいれば，交代化をおこなう順序を入れ換えても，結果が変わらないことを意味している．すなわち，2度目におこなう交代化を⟦…⟧で表すと，[…⟦…⟧…]=⟦…[…]…⟧である．もちろん，範囲が交わらなければ，[…]…⟦…⟧=⟦…⟧…[…]である．しかし，そのどちらでもないときには，一般に[…⟦…]…⟧≠⟦…[…⟧…]である．

交代化については公式

(vi)　$$rF_{[i_1\cdots i_r]}=\sum_{k=1}^{r}(-1)^{k-1}F_{i_k[i_1 i_2\cdots i_{k-1}i_{k+1}\cdots i_r]}$$
$$=F_{i_1[i_2\cdots i_r]}-\sum_{k=2}^{r}F_{i_k[i_2\cdots i_{k-1}i_1 i_{k+1}\cdots i_r]} \tag{2.5.10}$$

がしばしば有用である．(これを証明するには，$r!F_{[i_1\cdots i_r]}$ を定義通りに書き下して，$r!$個の項をFの先頭の添字に関してr個の類に分類し，符号を調べることによって，各類が(2.5.10)の中辺の各項を$(r-1)!$倍したものになっていることを示せばよい．)

さらに，r個のもののあらゆる順列($r!$通りある)を作るには，“まずr個からs個を取り出し($\binom{r}{s}$通りある)，取り出されたs個のあらゆる順列($s!$通りある)と残された$r-s$個のあらゆる順列($(r-s)!$通りある)を作る”というやり方があることに注意すると，ある固定した$s(<r)$について，

(vii)　$$r!F_{[i_1\cdots i_r]}=s!(r-s)!\sum\pm F_{[j_1\cdots j_s][j_{s+1}\cdots j_r]} \tag{2.5.11}$$

という形の関係が成り立つはずである($\sum$は$\{i_1,\cdots,i_r\}$の中からs個を選ぶあらゆる組合せ$\{j_1,\cdots,j_s\}$について取り，$\{j_1,\cdots,j_s,j_{s+1},\cdots,j_r\}=\{i_1,\cdots,i_r\}$とする)．符号を確定させるため，$i_1<\cdots<i_r, j_1<\cdots<j_s, j_{s+1}<\cdots<j_r$というように便宜的に添字の順を定める．$j_h=i_{k_h}(h=1,\cdots,r)$であるとして，(2.5.11)の右辺の各項の符号は，$(i_1\cdots i_r)$を$(j_1\cdots j_s j_{s+1}\cdots j_r)$に移すのに要する互換の回数を数えると，

$$(-1)^{k_1+\cdots+k_s+s(s+1)/2}=(-1)^{k_{s+1}+\cdots+k_r+(r-s)(r+s+1)/2} \tag{2.5.12}$$

であることもわかる．(2.5.10)はもちろん(2.5.11)の特別な場合である．

2.6 行　列　式

本節で述べる行列式の定義や諸性質は，理論的には重要であるが，実際に数値計算をするときには役に立たないものが多いことをあらかじめ注意しておく．

2.6.1 行列式の定義と基本性質

n 次正方行列 $A=[a_{ij}]$ の**行列式**(determinant)を

$$\det A=\sum \varepsilon^{1\,2\,\cdots\,n}_{j_1 j_2 \cdots j_n} a_{1j_1} a_{2j_2} \cdots a_{nj_n} \tag{2.6.1}$$

($\sum$ は $(1\cdots n)$ の $n!$ 個のすべての置換 $(j_1\cdots j_n)$ にわたって取る)で定義する．2.5 節の交代化演算の記号を用いれば(a_{ij} を a^i_j と書く)

$$\det A=n!\,a^1_{[1}a^2_2\cdots a^n_{n]}=n!\,a^{[1}_1 a^2_2\cdots a^{n]}_n$$
$$=n!\,a^{[1}_{[1}a^2_2\cdots a^{n]}_{n]} \tag{2.6.2}$$

とも表せることが直ちにわかる(交代化の性質(v)，(2.5.9)を見よ)．

注意 2.6.1　(2.6.2)の形からも明らかなように，A の行列式 $\det A$ の定義は，A の行と列の役割を入れ換えても変わらない．そこで，

$$\det A^{\mathrm{T}}=\det A \tag{2.6.3}$$

であるし，さらに，行列式のもつ諸性質についても，列に関して成り立つことは行に関しても成り立つ．

行列 A を列ベクトルの集まり $A=[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot n}]$ とみると(→2.4.1 項)，次の性質(i)〜(iv)が行列式の定義(2.6.1)(あるいは(2.6.2))から直ちに導かれる．

(i)　任意の j $(=1,\cdots,n)$ と $\lambda(\in K)$ に対して，

$$\det[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot j-1},\lambda\boldsymbol{a}_{\cdot j},\boldsymbol{a}_{\cdot j+1},\cdots,\boldsymbol{a}_{\cdot n}]$$
$$=\lambda\det[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot j-1},\boldsymbol{a}_{\cdot j},\boldsymbol{a}_{\cdot j+1},\cdots,\boldsymbol{a}_{\cdot n}]. \tag{2.6.4}$$

(ii)　任意の j とある n 次元縦ベクトル $\boldsymbol{b}$ に対して，

$$\det[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot j-1},\boldsymbol{a}_{\cdot j}+\boldsymbol{b},\boldsymbol{a}_{\cdot j+1},\cdots,\boldsymbol{a}_{\cdot n}]$$
$$=\det[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot j-1},\boldsymbol{a}_{\cdot j},\boldsymbol{a}_{\cdot j+1},\cdots,\boldsymbol{a}_{\cdot n}]$$
$$+\det[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot j-1},\boldsymbol{b},\boldsymbol{a}_{\cdot j+1},\cdots,\boldsymbol{a}_{\cdot n}]. \tag{2.6.5}$$

(iii)　列の交換により符号が反転する：

$$\det[\overset{1}{\boldsymbol{a}_{\cdot 1}},\cdots,\overset{j}{\boldsymbol{a}_{\cdot j}},\cdots,\overset{k}{\boldsymbol{a}_{\cdot k}},\cdots,\overset{n}{\boldsymbol{a}_{\cdot n}}]$$
$$=-\det[\overset{1}{\boldsymbol{a}_{\cdot 1}},\cdots,\overset{j}{\boldsymbol{a}_{\cdot k}},\cdots,\overset{k}{\boldsymbol{a}_{\cdot j}},\cdots,\overset{n}{\boldsymbol{a}_{\cdot n}}]. \qquad (2.6.6)$$

(iv)　単位行列 I_n に対して，

$$\det I_n=1. \qquad (2.6.7)$$

注意 2.6.2　上記の諸性質から，行列式を扱うとき有用な次の性質(イ)，(ロ)，(ハ)が導かれる．

(イ)　ある列の要素がみな 0 なら行列式の値は 0 である(性質(i)において $\lambda=0$ とおく；$\mathbf{0}$ は n 次元零縦ベクトル)：

$$\det[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot j-1},\mathbf{0},\boldsymbol{a}_{\cdot j+1},\cdots,\boldsymbol{a}_{\cdot n}]=0. \qquad (2.6.8)$$

(ロ)　二つ(以上)の等しい列があれば行列式の値は 0 である(性質(iii)による；体 K の標数が 2 であるときには，(iii)を経由しないで直接示すことができる)：

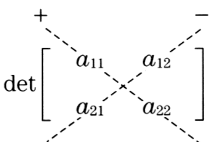
$$\det[\cdots,\boldsymbol{a},\cdots,\boldsymbol{a},\cdots]=0. \qquad (2.6.9)$$

(ハ)　列 j を $\lambda(\in K)$ 倍して列 k に加えても行列式の値は変わらない（性質(ii)，(i)，(ロ)をこの順で用いる)：

$$\det[\cdots,\boldsymbol{a}_{\cdot j},\cdots,\boldsymbol{a}_{\cdot k},\cdots]=\det[\cdots,\boldsymbol{a}_{\cdot j},\cdots,\boldsymbol{a}_{\cdot k}+\lambda\boldsymbol{a}_{\cdot j},\cdots]. \qquad (2.6.10)$$

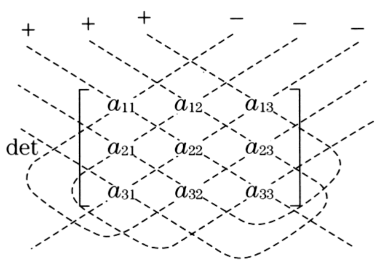
注意 2.6.1 でも注意したように，性質(i)～(iv)，(イ)～(ハ)は“列について”述べられているが，これらは“行について”も成り立つ．

2 次正方行列の行列式は，定義より直ちに，



$$\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}a_{22}-a_{12}a_{21} \qquad (2.6.11)$$

と表せる．(記憶の便宜上，破線とそれに符号を添記してある．) 3 次の場合も



$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33}+a_{12}a_{23}a_{31}+a_{13}a_{21}a_{32} -a_{11}a_{23}a_{32}-a_{12}a_{21}a_{33}-a_{13}a_{22}a_{31} \qquad (2.6.12)$$

と表せる．(破線とそれに添記した符号の役割は上と同様．)

しかし，4 次以上の場合にはこのような形には表せないことに注意されたい (→2.8 節)．

性質(iv)と(i)とから(定義からも直ちに)，対角行列の行列式はその対角要素の積に等しいことがわかる：

$$\det(\mathrm{diag}[\lambda_1,\cdots,\lambda_n])=\lambda_1\cdot\cdots\cdot\lambda_n. \tag{2.6.13}$$

置換行列の行列式は，性質(iii)と(iv)とから，1 あるいは -1 に等しい．特に互換行列の行列式は -1 である：

$$\det P(i_0,j_0)=-1. \tag{2.6.14}$$

基本行列 $T(i_0;t)$ については，性質(i)より，

$$\det T(i_0;t)=t \tag{2.6.15}$$

となり，$W(i_0,j_0;t)$ については，性質(ハ)より，

$$\det W(i_0,j_0;t)\,(=\det I)=1 \tag{2.6.16}$$

となる．

$W(i_0,j_0;t)$ は特殊な三角行列であるが，一般の三角行列の行列式は対角要素の積に等しい．たとえば，下三角行列 $L=[l_{ij}]$, $l_{ij}=0\,(i<j)$ については，

$$\det L=l_{11}\cdot\cdots\cdot l_{nn} \tag{2.6.17}$$

である．(2.6.17) は直接行列式の定義 (2.6.1) から導くのが早道である．$\sum$ の各項で $j_1=1$ 以外の(すなわち $j_1>1$ の)ものは $l_{1j_1}=0$ であるから 0 となる．$j_1=1$ の項では $j_2\neq j_1=1$ であるから $j_2\geqq 2$．しかし $j_2>2$ であると $l_{2j_2}=0$ となるので，0 でない可能性のある項では $j_2=2$．同様にして，$j_3=3,\cdots,j_n=n$ 以外の項が 0 となるので，残るのは項 $l_{11}l_{22}\cdots l_{nn}$ だけである(これも l_{ii} の一つが 0 ならば 0)．

$(n_1,\cdots,n_r;n_1,\cdots,n_r)$ という対称な分割の型をもつブロック下三角行列（上三角の場合も同様)

$$L=\begin{bmatrix} L_{11} & & & \\ L_{21} & L_{22} & & O \\ \vdots & \vdots & \ddots & \\ L_{r1} & L_{r2} & \cdots & L_{rr} \end{bmatrix} \tag{2.6.18}$$

に対しては，

$$\det L=(\det L_{11})\cdot(\det L_{22})\cdot\cdots\cdot(\det L_{rr}) \tag{2.6.19}$$

である．この公式も，行 i が第 k ブロックにあるとき$\left(i\leqq\sum_{h=1}^{k}n_h\right.$ であるが$\left.\right)$，$j>\sum_{h=1}^{k}n_h$ ならば，$l_{ij}=0$ であるということに着目すれば導くことができる．すなわち，(2.6.1)において $j_1,\cdots,j_{n_1}\leqq n_1$ でない項は 0 となるから，そして $j_1,\cdots,j_{n_1}$ は相異なるから，0 でない可能性のある項においては$\{j_1,\cdots,j_{n_1}\}=\{1,\cdots,n_1\}$である．$j_{n_1+1},\cdots,j_{n_1+n_2}$ は $j_1,\cdots,j_{n_1}$ と異なり，かつ $\leqq n_1+n_2$ でなければならないから，$\{j_{n_1+1},\cdots,j_{n_1+n_2}\}=\{n_1+1,\cdots,n_1+n_2\}$である．…．そのような $j_1,\cdots,j_n$ については，$N_k=\sum_{h=1}^{k-1}n_h$ とおくと($N_1=0$ と定義する)，

$$\varepsilon_{j_1\cdots j_n}^{1\cdots n}=\prod_{k=1}^{r}\varepsilon_{j_{N_k+1},\cdots,j_{N_k+n_k}}^{N_k+1,\cdots,N_k+n_k} \tag{2.6.20}$$

であるから，

$$\det L=\prod_{k=1}^{r}\left(\sum_{j_{N_k+1}=N_k+1}^{N_k+n_k}\cdots\sum_{j_{N_k+n_k}=N_k+1}^{N_k+n_k}\varepsilon_{j_{N_k+1},\cdots,j_{N_k+n_k}}^{N_k+1,\cdots,N_k+n_k}a_{N_k+1,j_{N_k+1}}\cdots a_{N_k+n_k,j_{N_k+n_k}}\right) \tag{2.6.21}$$

である．(2.6.21)の括弧の中は $\det L_{kk}$ に等しい．

注意 2.6.3 行列 A の行列式 $\det A$ は，A の要素の関数として，注意 2.6.1 の直後に述べた性質(i)～(iv)により特徴づける(定義する)こともできる．すなわち，

$$\boldsymbol{a}_{\cdot j}=\sum_{i=1}^{n}a_{ij}\boldsymbol{e}_i \tag{2.6.22}$$

と書けるので($\boldsymbol{e}_i$ は行 i の要素のみが 1 で他の要素がすべて 0 の縦ベクトル)，A の(要素の)関数 $\varphi(A)$ が性質(i)～(iv)を有するとすると，性質(ii)と(i)を繰り返し用いることによって，

$$\varphi(A)=\sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n}a_{i_1 1}\cdots a_{i_n n}\,\varphi([\boldsymbol{e}_{i_1},\cdots,\boldsymbol{e}_{i_n}]) \tag{2.6.23}$$

を得る．性質(iii)により(2.6.23)は，

$$\varphi(A)=\sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n}\varepsilon_{i_1\cdots i_n}^{1\cdots n}a_{i_1 1}\cdots a_{i_n n}\,\varphi([\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]) \tag{2.6.24}$$

と書き直せるが，$\varphi([\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n])=\varphi(I)$ は性質(iv)により $=1$ である．(2.6.24)の $\varphi(A)$ は，$\det A$ の定義式(2.6.1)(の行と列の役割を入れ換えたもの，すなわち(2.6.2)の第 2 式)に他ならない．

2.6.2 行列式の既約性

n 次正方行列 A の要素 a_{ij} をすべて独立な変数とみなせば，$\det A$ は n^2 個の変数の n 次多項式である．そのような多項式として $\det A$ は既約であることが

知られている．実際，もし $\det A=\varphi(A)\psi(A)$（$\varphi(A)$ も $\psi(A)$ も n^2 個の変数 $a_{11},\cdots,a_{nn}$ の多項式）と積の形に分解されたとすると，

- $\det A$ は $a_{1j}(j=1,\cdots,n)$ に関して 1 次であり（すなわち，それらの積の形を含まない），しかもそれらすべてに実際に依存するから，それらの変数は φ か ψ の一方だけに実際に含まれる（φ に含まれるとする）；
- また，$\det A$ は $a_{1j}a_{ij}(i\neq 1)$ という形の積を含まないから，a_{1j} が $\varphi(A)$ に含まれれば $a_{ij}(i\neq 1)$ は $\psi(A)$ には含まれないで，$\varphi(A)$ の方に含まれる．

このようにして，$\psi(A)$ は a_{ij} のどれも含まない定数でしかありえない，すなわち，$\det A$ は既約である．

注意 2.6.4　$\det A$ の A の要素の多項式としての既約性は，A が対称行列でその上三角部分の要素だけが独立な変数であるとした場合でも成り立っている．（上の証明を少し変更すればよい．）しかし，偶数次の反対称行列 A については，後に (2.9.1) に示すように，$\det A=(\varphi(A))^2$（$\varphi(A)$ は $a_{ij}(i<j)$ の多項式）という形になっている．（奇数次の反対称行列の行列式は次に示すように恒等的に 0 である．）

2.6.3　特殊な形の行列の行列式

性質 (i) により，n 次正方行列 A に対して $\det(\lambda A)=\lambda^n \det A$ であるが，特に A が反対称，すなわち，$A^{\mathsf{T}}=-A$ であれば，$\det A=\det A^{\mathsf{T}}=\det(-A)=(-1)^n \det A$．$A$ が奇数次の反対称行列であれば（K の標数が 2 でないとして），$\det A=0$ である．偶数次の反対称行列の行列式については 2.9 節の Pfaffian の項を見よ．K の標数が 2 のときでも，奇数次の反対称行列（これは対角要素が 0 の対称行列であるといっても同じ）の行列式は 0 であることも示せる（→2.15.4 項 c）．より詳しくは，後に 2.15.4 項 c で示されるように，一般の反対称行列の階数（→2.13.1 項）は偶数である．

同様に，$K=\mathbb{C}$ のとき，Hermite 行列の行列式は実数（$\det A=\det A^{\mathsf{H}}=\overline{\det A}$ より），偶数次の歪 Hermite 行列の行列式は実数，奇数次の歪 Hermite 行列の行列式は純虚数である（$\det A=(-1)^n \det A^{\mathsf{H}}=(-1)^n\overline{\det A}$ より）．

2.7 小行列式，余因子，余因子行列

2.7.1 小行列式

行列 A(正方とは限らない)の正方部分行列 $A^{i_1\cdots i_r}_{j_1\cdots j_r}$ の行列式 $\det A^{i_1\cdots i_r}_{j_1\cdots j_r}$ を A の r 次の**小行列式**(subdeterminant, minor)といい，$|A|^{i_1\cdots i_r}_{j_1\cdots j_r}$ と記すこともある．(特に，$|A|^{i}_{j}=a_{ij}$ である．) 以降では，$i_1,\cdots,i_r$ の中に重複があったり $j_1,\cdots,j_r$ の中に重複があったりしてもよいとするが，そのような場合には小行列式の値は 0 であるとする．

正方行列 A の小行列式で $|A|^{i_1\cdots i_r}_{i_1\cdots i_r}$ の形のものを**主小行列式**(principal minor)，$|A|^{12\cdots r}_{12\cdots r}$ の形のものを**首座小行列式**(leading principal minor)と呼ぶことがある．

$m\times n$ 型の行列 A に対して，そのすべての r 次小行列式 $|A|^{i_1\cdots i_r}_{j_1\cdots j_r}$ を考える($r\leqq\min(m,n)$)．無意味なもの(行や列に重複があって 0 になるもの，行や列の順序だけが異なるもの)を除くために，便宜上，$1\leqq i_1<i_2<\cdots<i_r\leqq m$，$1\leqq j_1<j_2<\cdots<j_r\leqq n$ を満たすように添字を並べることにする．このような小行列式は全部で$\binom{m}{r}\cdot\binom{n}{r}$個あるが，それらを行方向には$(i_1,\cdots,i_r)$の辞書式順に並べ，列方向には$(j_1,\cdots,j_r)$の辞書式順に並べてできる$\binom{m}{r}\times\binom{n}{r}$型の行列を $A^{<r>}$ と表すことにする．(特に，$A^{<1>}=A$ である．) A が正方行列なら，$A^{<r>}$ も正方で $A^{<r>}$ の対角要素が A の r 次の主小行列式である．

2.7.2 トレース

n 次正方行列 $A=[a_{ij}]$ に対して，**トレース**(**跡**ともいう；trace)[1] $\mathrm{tr}\,A$ を

$$\mathrm{tr}\,A=\sum_{i=1}^{n}a_{ii} \tag{2.7.1}$$

で定義する．A,B が共に n 次正方行列のとき，

$$\begin{aligned}&\mathrm{tr}(A+B)=\mathrm{tr}\,A+\mathrm{tr}\,B,\\&\mathrm{tr}(A^{\mathsf{T}})=\mathrm{tr}\,A,\qquad \mathrm{tr}(\lambda A)=\lambda\,\mathrm{tr}\,A,\\&\mathrm{tr}(\bar{A})=\mathrm{tr}(A^{\mathsf{H}})=\overline{(\mathrm{tr}\,A)}\qquad (K=\mathbb{C}\text{ のとき})\end{aligned} \tag{2.7.2}$$

などは自明であるが，A が $m\times n$ 型で B が $n\times m$ 型のとき，

1) $\mathrm{tr}\,A$ は $\mathrm{Spur}\,A$ と書いて行列 A のシュプールと呼んでいたことがある．これは英語の trace に相当するドイツ語が Spur であるからである．

$$\mathrm{tr}(AB)=\mathrm{tr}(BA)=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}b_{ji} \tag{2.7.3}$$

という関係も，容易に導き出せて，有用である．(2.7.3)を一般化した

$$\mathrm{tr}(ABC)=\mathrm{tr}(BCA)=\mathrm{tr}(CAB)=\sum_{i=1}^{l}\sum_{j=1}^{m}\sum_{k=1}^{n}a_{ij}b_{jk}c_{ki} \tag{2.7.4}$$

などもある(A は $l\times m$ 型，B は $m\times n$ 型，C は $n\times l$ 型；4個以上の積についても同様)．

A が m 次正方行列，B が n 次正方行列のとき，Kronecker 積(2.4.2項)のトレースは

$$\mathrm{tr}(A\otimes B)=(\mathrm{tr}\,A)\cdot(\mathrm{tr}\,B)\,. \tag{2.7.5}$$

n 次正方行列 A に対して，その r 次の主小行列式の和 $\mathrm{tr}\,A^{<r>}$ ($r=1,2,\cdots,n$) もいろいろな場面で重要な役割を果たす．(特に，$\mathrm{tr}\,A^{<1>}=\mathrm{tr}\,A$，$\mathrm{tr}\,A^{<n>}=\det A$ である．)

注意 2.7.1　$\mathrm{tr}\,A^{<r>}$ は c_r と記すことも多いが，交代化演算の記号を用いると，

$$\sum_{\substack{\{i_1,\cdots,i_r\}\\ \subseteq\{1,\cdots,n\}}} a^{[i_1}_{i_1}\cdots a^{i_r]}_{i_r}=\sum_{\substack{\{i_1,\cdots,i_r\}\\ \subseteq\{1,\cdots,n\}}}|A|^{i_1\cdots i_r}_{i_1\cdots i_r}=c_r \tag{2.7.6}$$

と書ける．

2.7.3　余因子，余因子行列

n 次正方行列 A の (i,j) 要素 a_{ij} の**余因子**(cofactor)は，

$$\Delta_{ij}=(-1)^{i+j}|A|^{1,\cdots\cdots,i-1,i+1,\cdots,n}_{1,\cdots,j-1,j+1,\cdots\cdots,n} \tag{2.7.7}$$

で定義される(行列 A の行 i と列 j を消してできる $n-1$ 次正方行列の行列式に符号 $(-1)^{i+j}$ をつけたもの)．A の部分行列 $A^{i_1\cdots i_r}_{j_1\cdots j_r}$ ($i_1<i_2<\cdots<i_r$，$j_1<j_2<\cdots<j_r$ とする)の余因子は，

$$\Delta^{i_1\cdots i_r}_{j_1\cdots j_r}=(-1)^{i_1+\cdots+i_r+j_1+\cdots+j_r}|A|^{i_{r+1}\cdots i_n}_{j_{r+1}\cdots j_n} \tag{2.7.8}$$

で定義される．ここで，$i_{r+1},\cdots,i_n$ は $i_{r+1}<\cdots<i_n$，$\{i_1,\cdots,i_r,i_{r+1},\cdots,i_n\}=\{1,\cdots,n\}$ を満たすように定める(一意に定まる)．$j_{r+1},\cdots,j_n$ についても同様．

A の**余因子行列**(**随伴行列**；adjoint matrix，adjugate matrix) $\hat{A}=[\hat{a}_{ij}]$ は，

$$\hat{a}_{ij}=\Delta_{ji}\qquad(i,j\ \text{の順番に注意}) \tag{2.7.9}$$

で定義する．$\hat{A}$ は adj A と書かれることが多い．さらに，A からその r 次の小行列式の行列 $A^{<r>}$ を作ったのにならって，$\Delta^{i_1\cdots i_r}_{j_1\cdots j_r}$ を要素とする $\binom{n}{r}\times\binom{n}{r}$ 型の行

列 $\tilde{A}^{<r>}$ を作る．ただし，$\Delta^{i_1\cdots i_r}_{j_1\cdots j_r}$ の $j_1,\cdots,j_r$ が $\tilde{A}^{<r>}$ の行(列ではない)を指し，$i_1,\cdots,i_r$ が $\tilde{A}^{<r>}$ の列(行ではない)を指すものとする．(特に，$\tilde{A}^{<1>}=\hat{A}$ である．) また，$\tilde{A}^{<n-r>}$ は $(A^{<r>})^{\top}$ と，行および列の並び順と符号を除いて等しい．特に，$\hat{A}$ は $(A^{<n-1>})^{\top}$ とそのような関係にある．

2.8　行列式の展開

2.8.1　Laplace 展開

n 次正方行列 A の行列式の定義(2.6.1)と交代化演算による表現(2.6.2)に交代化演算の性質(特に(2.5.10))を適用し，それに余因子行列の定義(2.7.7)，(2.7.9)を用いると，任意の i_0 に対して，

$$\begin{aligned}\det A &= n!\,a^1_{[1}a^2_2\cdots a^n_{n]} = (-1)^{i_0-1}n!\,a^{i_0}_{[1}a^1_2\cdots a^{i_0-1}_{i_0}a^{i_0+1}_{i_0+1}\cdots a^n_{n]}\\ &= (-1)^{i_0-1}(n-1)!\sum_{j=1}^{n}(-1)^{j-1}a^{i_0}_j a^1_{[1}\cdots{}^{\check{i_0}}\cdots{}_{\hat{j}}\cdots a^n_{n]}\\ &= \sum_{j=1}^{n}a^{i_0}_j(-1)^{i_0+j}|A|^{1\cdots\check{i_0}\cdots n}_{1\cdots\hat{j}\cdots n}\\ &= \sum_{j=1}^{n}a_{i_0j}\Delta_{i_0j} = \sum_{j=1}^{n}a_{i_0j}\hat{a}_{ji_0} \qquad (2.8.1)\end{aligned}$$

が得られる．($\check{i_0}$ や $\hat{j}$ は，連続した添字番号がそこだけとんでいることを示す．)

(2.8.1)は“$\det A$ は A の行 i_0 の各要素にその余因子を掛けて加えたものに等しい”と読める．(i_0 は $1,\cdots,n$ のどれでもよい．) これを，行列式 $\det A$ の行 i_0 に関する **Laplace(ラプラス)展開**(Laplace expansion)という．

同様にして，列 j_0 に関する Laplace 展開

$$\det A = \sum_{i=1}^{n}\Delta_{ij_0}a_{ij_0} = \sum_{i=1}^{n}\hat{a}_{j_0i}a_{ij_0} \qquad (2.8.2)$$

も得られる．

Laplace 展開は n 次正方行列の行列式の計算を $n-1$ 次の行列式の計算に帰着させることであるともみなせる．ただし，これを繰り返し適用して行列式の次数を次々と減らしていけば，結局，行列式の定義式(2.6.1)に到達してしまう．

(2.8.1)や(2.8.2)は行列 A とその余因子行列 $\hat{A}$ との積 $A\hat{A}$ あるいは $\hat{A}A$ の対角要素((i_0,i_0) 要素あるいは (j_0,j_0) 要素)が $\det A$ に等しいということを表しているとみなせる．それでは $A\hat{A}$ や $\hat{A}A$ の非対角要素はどうなるであろうか．た

とえば，(2.8.1)において，

$$\sum_{j=1}^{n} a_{i_0 j}\hat{a}_{j i_1} \qquad (i_0 \neq i_1) \tag{2.8.3}$$

としてみる．これは式中で $\varDelta_{i_0 j}$ の代りに $\varDelta_{i_1 j}$ を用いることを意味する．ところで，A の行 i_1 のところに行 i_0 の要素と同じものをおいた行列を A' としてみる(行 i_0 自身はそのまま)と，A' に対する $\varDelta'_{i_1 j}$ は $\varDelta'_{i_1 j}=\pm\varDelta_{i_0 j}$ ($\pm$ は j によらない)であることは明らかであるから，(2.8.3)は det A' の行 i_1 に関する Laplace 展開に，符号を除いて，一致している．A' は行 i_0 と i_1 が等しいから，det $A'=0$ である．このようにして，(2.8.3)は 0 になることが示された．以上をまとめると，

$$A\hat{A}=(\det A)\cdot I_n \tag{2.8.4}$$

ということになる．列に関する Laplace 展開について考えれば，

$$\hat{A}A=(\det A)\cdot I_n \tag{2.8.5}$$

も成り立つことがわかる．

2.8.2 一般化 Laplace 展開

Laplace 展開は複数行に関する形に一般化できる．すなわち，n 次正方行列 A の行 $1,\cdots,n$ のある与えられた部分集合 $i_1,\cdots,i_r$ ($i_1<\cdots<i_r$ とする)に関して，

$$\det A=\sum |A|^{i_1\cdots i_r}_{j_1\cdots j_r}\varDelta^{i_1\cdots i_r}_{j_1\cdots j_r}=\sum \pm |A|^{i_1\cdots i_r}_{j_1\cdots j_r}|A|^{i_{r+1}\cdots i_n}_{j_{r+1}\cdots j_n} \tag{2.8.6}$$

である．($\sum$ は $j_1<\cdots<j_r$ であるようなすべての $j_1,\cdots,j_r$ の組合せ($\binom{n}{r}$個ある)について取る．$i_{r+1},\cdots,i_n,j_{r+1},\cdots,j_n$，および符号$\pm$に関しては(2.7.8)を見よ．)

あるいは，A の r 次小行列式を要素とする行列 $A^{\langle r\rangle}$ と $n-r$ 次の余因子を要素とする行列 $\tilde{A}^{\langle r\rangle}$ (→2.7 節)を用いて表せば，

$$A^{\langle r\rangle}\tilde{A}^{\langle r\rangle}=\tilde{A}^{\langle r\rangle}A^{\langle r\rangle}=(\det A)\cdot I_{\binom{n}{r}} \tag{2.8.7}$$

が成り立つ．実際，$\{i_{r+1},\cdots,i_n\}=\{1,\cdots,n\}-\{i_1,\cdots,i_r\}$ で $i_{r+1}<\cdots<i_n$ として，(2.6.2)をまず

$$\det A=n!\,a^{[1}_{1}\cdots a^{n]}_{n}=(-1)^{i_1+\cdots+i_r+r(r+1)/2}n!\,a^{[i_1}_{1}\cdots a^{i_r}_{r}a^{i_{r+1}}_{r+1}\cdots a^{i_n]}_{n} \tag{2.8.8}$$

と書き直してから(符号は，$1,\cdots,n$ の並びから出発して，i_1 を左端に移動させるのに互換を i_1-1 回，次に i_2 を i_1 の右隣にまで移動させるのに i_2-2 回，…，それぞれ実行すればよい——$i_{r+1},\cdots,i_n$ は自然にこの順に並んで残る——ということに注意すれば決定される)，(2.5.11)(符号は(2.5.12))を適用すると

$$\det A = n!\,\varepsilon_1 a^{[i_1}_{1}\cdots a^{i_r}_{r} a^{i_{r+1}}_{r+1}\cdots a^{i_n]}_{n} = n!\,\varepsilon_1 a^{i_1}_{[1}\cdots a^{i_r}_{r} a^{i_{r+1}}_{r+1}\cdots a^{i_n}_{n]}$$
$$= r!\,(n-r)!\,\varepsilon_1 \sum \varepsilon_2 a^{i_1}_{[j_1}\cdots a^{i_r}_{j_r]} a^{i_{r+1}}_{[j_{r+1}}\cdots a^{i_n}_{j_n]} \qquad (2.8.9)$$

を得る．ここで，$\sum$ は$\{1,\cdots,n\}$を$\{j_1,\cdots,j_r\},\{j_{r+1},\cdots,j_n\}$($j_1<\cdots<j_r; j_{r+1}<\cdots<j_n$)に分けるあらゆる分け方について取る．また，符号 $\varepsilon_1, \varepsilon_2$ は，

$$\varepsilon_1 = (-1)^{i_1+\cdots+i_r+r(r+1)/2},$$
$$\varepsilon_2 = (-1)^{j_1+\cdots+j_r+r(r+1)/2}$$

より，

$$\varepsilon_1\varepsilon_2 = (-1)^{i_1+\cdots+i_r+j_1+\cdots+j_r} \qquad (2.8.10)$$

である．そこで，結局，(2.8.9)は(2.8.6)(すなわち(2.8.7))の対角要素の間の等式を表していることがわかる．

$\{i_1',\cdots,i_r'\}\neq\{i_1,\cdots,i_r\}$のときには，上と同様にして，(2.8.7)の左辺の要素は，

$$n!\,a^{[i_1'}_{1}\cdots a^{i_r'}_{r} a^{i_{r+1}}_{r+1}\cdots a^{i_n]}_{n} \qquad (2.8.11)$$

と表されるが，$\{i_1',\cdots,i_r'\}\neq\{i_1,\cdots,i_r\}$なら $i_1',\cdots,i_r'$ の中に $i_{r+1},\cdots,i_n$ のどれかと等しいものがあるので，(2.8.11)は上の添字の交代化によって 0 となる．これが(2.8.7)の非対角要素の間の等式を表している．(2.8.7)は(2.8.4)の一般化である．

(ブロック)三角行列の行列式に対する公式(2.6.17)，(2.6.19)は Laplace 展開の公式(2.8.4)，(2.8.7)を繰り返し適用することによっても容易に導くことができる．

注意 2.8.1　上の 2.8.1 項，2.8.2 項の Laplace 展開の公式の本質は交代化演算の性質 2.5.2 節の(2.5.10)，(2.5.11)にあることに注意されたい．

2.8.3 Binet-Cauchy 展開

A が $m\times n$ 型，B が $n\times m$ 型の行列のとき，

$$\det(AB) = m!\sum_{j_1=1}^{n}\cdots\sum_{j_m=1}^{n} a^{[1}_{j_1}\cdots a^{m]}_{j_m} b^{j_1}_{[1}\cdots b^{j_m}_{m]} \qquad (2.8.12)$$

であるから(交代化演算と $\sum$ とは可換であること，そして(2.6.2)に注意)，これに交代化演算の性質(2.5.5)を適用すると，

$$\det(AB) = m!\sum_{j_1=1}^{n}\cdots\sum_{j_m=1}^{n} a^{[1}_{[j_1}\cdots a^{m]}_{j_m]} b^{[j_1}_{1}\cdots b^{j_m]}_{m} \qquad (2.8.13)$$

となる．したがって，$j_1,\cdots,j_m$ の中に等しいものが重複して現れるような項は $\sum\cdots\sum$ の中で消し合う．$\{1,\cdots,n\}$から異なる m 個のものを選んで $j_1<\cdots<j_m$ の順に並べたものを基準にすると，和の中に同類項が $m!$ 個ずつ現れる．また，符号は，同じ$[j_1,\cdots,j_m]$が a 側にも b 側にもあるから，どの項についても＋である．このようにして，

$$\det(AB)=\sum|A|^{1\cdots m}_{j_1\cdots j_m}|B|^{j_1\cdots j_m}_{1\cdots m} \tag{2.8.14}$$

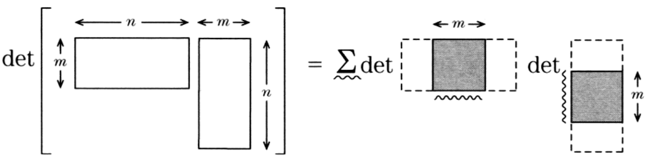
(和は$\{1,\cdots,n\}$から異なる m 個を取る組合せ$\{j_1,\cdots,j_m\}$($j_1<\cdots<j_m$ の順に並べる)すべてについて取る)が得られる．特に，$m>n$ ならば，上のような $j_1,\cdots,j_m$ は選べないから，

$$m>n \text{ のとき}\quad \det(AB)=0 \tag{2.8.15}$$

である．また，$m=n$ ならば，上のような j_k は $j_k=k\,(k=1,\cdots,m)$ しかないから，

$$A,B \text{ が共に正方行列のとき}\quad \det(AB)=(\det A)\cdot(\det B) \tag{2.8.16}$$

となる．以上の $\det(AB)$を表す公式を **Binet-Cauchy**(ビネー・コーシー)**の公式**という．直観的なイメージで述べると，



(2.8.14′)

($\sum$ は連続していない〰〰についても取る)である．

AB の小行列式についても同様の議論が成り立つので，2.7.1項の記号を用いて，一般に，

$$(AB)^{<r>}=A^{<r>}\cdot B^{<r>} \tag{2.8.17}$$

という関係がある．(この場合，A の列の数と B の行の数が等しければ，AB は正方行列でなくともよい．)

これを繰り返し適用すれば，同様にして，

$$(ABC\cdots)^{<r>}=A^{<r>}B^{<r>}C^{<r>}\cdots \tag{2.8.18}$$

が得られる．

2.9 Pfaffian

A が奇数次の反対称行列のとき，$\det A=0$ であることは既に述べたが(→2.6節)，偶数次($n=2m$ 次)の反対称行列のときには，$\det A$ は

$$\det A=\left(\frac{n!}{\left(\frac{n}{2}\right)!2^{n/2}}a_{[12}a_{34}\cdots a_{n-1,n]}\right)^2 \tag{2.9.1}$$

と表せる．(2.9.1)の括弧の中

$$\mathrm{Pf}\,A=\frac{n!}{\left(\frac{n}{2}\right)!2^{n/2}}a_{[12}a_{34}\cdots a_{n-1,n]} \tag{2.9.2}$$

は A の **Pfaffian**(パフィアン，プファフィアン)と呼ばれる．

A が反対称 $a_{ij}=-a_{ji}$ であるから，a_{ij} のうち $i<j$ であるものを a_{ij} と a_{ji} の代表に選ぶことにしてすべての式を書くことにする．すると，(2.9.2)は $\{1,\cdots,n=2m\}$を m 個の“対”$(i_1,j_1),\cdots,(i_m,j_m)$(ただし，$i_k<j_k$，$i_1<\cdots<i_m$ とする；この約束から常に $i_1=1$，$i_m\leqq n-1$ となる)に分ける分け方すべてについて，

$$\varepsilon^{1\ 2\ \cdots\ n-1\ n}_{i_1 j_1\cdots\ i_m\ j_m}a_{i_1j_1}\cdots a_{i_mj_m} \tag{2.9.3}$$

を加え合わせたものということになる．(2.9.2)の右辺の一見面倒な係数は交代化の結果現れる同類項が1回だけ数えられるように定めているにすぎない．

たとえば

$$A=\begin{bmatrix} 0 & a_{12} & a_{13} & a_{14} \\ -a_{12} & 0 & a_{23} & a_{24} \\ -a_{13} & -a_{23} & 0 & a_{34} \\ -a_{14} & -a_{24} & -a_{34} & 0 \end{bmatrix}$$

に対して，$\mathrm{Pf}\,A$ を定義通りに，$\det A$ を Laplace 展開などで計算すると，

$$\begin{aligned}\mathrm{Pf}\,A&=a_{12}a_{34}-a_{13}a_{24}+a_{14}a_{23},\\ \det A&=(a_{12}a_{34})^2+(a_{13}a_{24})^2+(a_{14}a_{23})^2\\ &\quad-2a_{12}a_{13}a_{24}a_{34}+2a_{12}a_{14}a_{23}a_{34}-2a_{13}a_{14}a_{23}a_{24}\\ &=(\mathrm{Pf}\,A)^2\end{aligned}$$

となる．(2.9.1)が成立することは，純粋に組合せ論的議論によって示すこともできるし，その方がより本質的な構造が明らかになるのであるが，それは補章A.5節に譲り，本章の中では2.15.4項cにおいて別のやり方で示すことにする．

注意2.9.1　n次正方行列Aの行列式$\det A=|A|_{1\cdots n}^{1\cdots n}$に対して小行列式$|A|_{j_1\cdots j_r}^{i_1\cdots i_r}$を考えたように(2.7.1項)，$n$次反対称行列$A$のPfaffianに対して**小Pfaffian**(subpfaffian)とでもいうべきものを

$$(\mathrm{Pf}\,A)_{i_1\cdots i_{2q}}=\frac{(2q)!}{q!2^q}a_{[i_1i_2}a_{i_3i_4}\cdots a_{i_{2q-1}i_{2q}]} \tag{2.9.4}$$

で定義することができる($i_1,i_2,\cdots,i_{2q}$は$\{1,\cdots,n\}$から選んだすべて異なる番号)．すると，階数(→2.13.1項)が$r=2q(\leqq n)$の反対称行列Aの$r=2q$次の小行列式(これは$r<n$のときには複数個あることに注意)を小Pfaffianで表す関係式

$$|A|_{j_1\cdots j_{2q}}^{i_1\cdots i_{2q}}=(\mathrm{Pf}\,A)_{i_1\cdots i_{2q}}(\mathrm{Pf}\,A)_{j_1\cdots j_{2q}} \tag{2.9.5}$$

が成立する．これについても2.15.4項cで示す．

注意2.9.2　Pfaffianと行列式との間には多くの興味深い関係が知られているが，それらについては補章A.5節で述べる．次の関係は，非常に簡単ではあるが，行列式よりもPfaffianの方がより基本的な概念ではないかと思わせるものである．

任意のn次正方行列$A=[a_{ij}]\,(i,j=1,\cdots,n)$に対して$2n$次の反対称行列

$$B=[b_{\tilde{\imath}\tilde{\jmath}}]=\begin{array}{c|c|c|}
 & 1\ \cdots\ n & n+1\cdots 2n \\ \hline
1,\cdots,n & O_{n,n} & A \\ \hline
n+1,\cdots,2n & -A^{\mathsf{T}} & O_{n,n} \\ \hline
\end{array} \tag{2.9.6}$$

を対応させる．ここで，

$$\tilde{\imath},\tilde{\jmath}=1,2,\cdots,n\,;\,n+1,\cdots,2n, \tag{2.9.7}$$

$$b_{\tilde{\imath}\tilde{\jmath}}=\begin{cases}0 & (1\leqq\tilde{\imath},\tilde{\jmath}\leqq n\text{ あるいは }n+1\leqq\tilde{\imath},\tilde{\jmath}\leqq 2n\text{ のとき}),\\ a_{\tilde{\imath},\tilde{\jmath}-n} & (1\leqq\tilde{\imath}\leqq n,\ n+1\leqq\tilde{\jmath}\leqq 2n\text{ のとき}),\\ -a_{\tilde{\jmath},\tilde{\imath}-n} & (n+1\leqq\tilde{\imath}\leqq 2n,\ 1\leqq\tilde{\jmath}\leqq n\text{ のとき})\end{cases} \tag{2.9.8}$$

である．Pf Bの展開の同類項の代表を上の約束に従って表したとき，0にならないもの$b_{\tilde{\imath}_1\tilde{\jmath}_1}\cdots b_{\tilde{\imath}_n\tilde{\jmath}_n}$において，$\tilde{\imath}_k<\tilde{\jmath}_k$と(2.9.8)の第1式とにより$1\leqq\tilde{\imath}_k\leqq n$でなければならず，また$\tilde{\imath}_1<\cdots<\tilde{\imath}_n$より$\tilde{\imath}_1=1$, $\tilde{\imath}_2=2,\cdots,\tilde{\imath}_n=n$でなければならない．同様に，$n+1\leqq\tilde{\jmath}_k\leqq 2n$であるから，$j_k=\tilde{\jmath}_k-n$とおけば，展開の代表項は

$$\pm a_{1j_1}a_{2j_2}\cdots a_{nj_n} \tag{2.9.9}$$

という形になる．この前につく符号は

$$\varepsilon^{1,2,\cdots,n,n+1,\cdots,2n}_{1,\tilde{\jmath}_1,\cdots\cdots\cdots\cdots,n,\tilde{\jmath}_n} \tag{2.9.10}$$

である．$(1\ \tilde{\jmath}_1\ 2\ \tilde{\jmath}_2\cdots n\ \tilde{\jmath}_n)$を$(1\ 2\cdots n\ \tilde{\jmath}_1\ \tilde{\jmath}_2\cdots\tilde{\jmath}_n)$に移動するのに要する互換の数は，

$2,3,\cdots,n$ を左に送って行くことを考えると，$(2-1)+(3-1)+\cdots+(n-1)=n(n-1)/2$ であるから，

$$\begin{aligned}\varepsilon^{1,2\cdots,n,n+1,\cdots,2n}_{1,\bar{j}_1,\cdots\cdots\cdots\cdots,n,\bar{j}_n}&=(-1)^{n(n-1)/2}\varepsilon^{1,\cdots,n,n+1,\cdots,2n}_{1,\cdots,n,\bar{j}_1,\cdots\cdots,\bar{j}_n}\\&=(-1)^{n(n-1)/2}\varepsilon^{n+1,\cdots,2n}_{\bar{j}_1,\cdots\cdots,\bar{j}_n}=(-1)^{n(n-1)/2}\varepsilon^{1\ \cdots n}_{j_1\cdots j_n}\end{aligned}\tag{2.9.11}$$

となり，これをすべての可能な異なる $j_1,\cdots,j_n\in\{1,\cdots,n\}$ について加え合わせると，結局行列式の定義(2.6.1)と同じ形になり

$$\det A=(-1)^{n(n-1)/2}\,\mathrm{Pf}\,B\tag{2.9.12}$$

が得られる．

注意 2.9.3 任意の $n=2m$ 次の正方行列 A(反対称であるとは限らない)に対して

$$B=ADA^{\mathsf{T}},$$

$$D=\begin{pmatrix}\begin{matrix}0&1\\-1&0\end{matrix}&&&O\\&\begin{matrix}0&1\\-1&0\end{matrix}&&\\&&\ddots&\\O&&&\begin{matrix}0&1\\-1&0\end{matrix}\end{pmatrix}\tag{2.9.13}$$

によって $2m$ 次の反対称行列 B を作る．($B^{\mathsf{T}}=(ADA^{\mathsf{T}})^{\mathsf{T}}=AD^{\mathsf{T}}A^{\mathsf{T}}$，$D^{\mathsf{T}}=-D$ であるから，$B^{\mathsf{T}}=-B$ で，B は反対称である．) このとき

$$\det A=\mathrm{Pf}\,B\tag{2.9.14}$$

であることも同様にして示すことができる．実際，記号の混乱を避けるために a_{ij} をあらためて a_i^j と記すことにすると[1]，

$$b_{ij}=\sum_{k=1}^{m}(a_i^{2k-1}a_j^{2k}-a_i^{2k}a_j^{2k-1})=2\sum_{k=1}^{m}a_{[i}^{2k-1}a_{j]}^{2k}\tag{2.9.15}$$

と表せる．そこで，

$$\begin{aligned}\mathrm{Pf}\,B&=\frac{(2m)!}{m!2^m}b_{[12}b_{34}\cdots b_{2m-1,2m]}\\&=\frac{(2m)!}{m!}\left(\sum_{k_1=1}^{m}a_{[1}^{2k_1-1}a_2^{2k_1}\right)\left(\sum_{k_2=1}^{m}a_3^{2k_2-1}a_4^{2k_2}\right)\cdots\left(\sum_{k_m=1}^{m}a_{2m-1}^{2k_m-1}a_{2m]}^{2k_m}\right)\end{aligned}$$

を計算してみる．右辺の括弧をほどいて得られる各項

$$(2m)!\,a_{[1}^{2k_1-1}a_2^{2k_1}a_3^{2k_2-1}a_4^{2k_2}\cdots a_{2m-1}^{2k_m-1}a_{2m]}^{2k_m}$$

は，$k_1,\cdots,k_m$ の中に等しいものがあるときには 0 となり(2.5.2 項の交代化の性質(i)，(2.5.5)による)，それらがすべて異なるときには

$$(2m)!\,a_{[1}^1a_2^2\cdots a_{2m-1}^{2m-1}a_{2m]}^{2m}$$

1) 他所では a_j^i と書くことが多いが，ここでは形を見やすくするために，記号を敢えて a_i^j としてあることに注意されたい．

に等しい．（対$(2k_i-1,2k_i)$と対$(2k_j-1,2k_j)$を入れ換える置換は偶置換であることに注意．）2.6節の行列式の定義(2.6.2)によりこれは$\det A$に等しい．このような項は$(1,\cdots,m)$のすべての順列$(k_1,\cdots,k_m)$について加え合わせられるから，さらに$m!$倍され，結局

$$\mathrm{Pf}\, B=\det A \tag{2.9.16}$$

と表されることになる．

なお，Aが奇数次の正方行列であっても，$A'=\left[\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array}\right]$を考えることによって，$\det A=\det A'$であるから，

$$B'=A'D'A'^{\mathrm{T}}$$

とおけば

$$\det A=\mathrm{Pf}\, B' \tag{2.9.17}$$

と表すことができる．

2.10 逆　行　列

2.10.1 正則行列と逆行列

n次正方行列Aの行列式$\det A$が0でないとき，

$$A^{-1}=(\det A)^{-1}\hat{A} \tag{2.10.1}$$

によりA^{-1}を定義すると，Laplace展開の式(2.8.4)，(2.8.5)により，

$$AA^{-1}=A^{-1}A=I \tag{2.10.2}$$

となる．A^{-1}をAの**逆行列**(inverse)と呼ぶ．

注意 2.10.1　A^{-1}は，行列環$\mathscr{M}_n$(注意2.2.4)において，乗法に関するAの逆元である．また，「もし，Aが逆転可能，すなわち$AB=I$なるBが存在するならば，(2.8.16)により$\det A\neq 0$であるから，$A^{-1}=A^{-1}I=A^{-1}(AB)=(A^{-1}A)B=B$である．ゆえに，$BA=I$でもあり，そして，このような$B$が一意に定まる．$BA=I$についても同様．」

$\det A\neq 0$ならA^{-1}が存在するが，逆に，Aに対して(2.10.2)を満たすA^{-1}が存在すれば，(2.8.16)により$(\det A)\cdot(\det A^{-1})=\det I=1$であり，したがって，$\det A\neq 0$，$\det A^{-1}\neq 0$，かつ

(i)　$\det A^{-1}=(\det A)^{-1}$　(2.10.3)

である．

逆行列をもつ(あるいは行列式が 0 でない)行列のことを**正則**(regular)あるいは**非特異**(nonsingular)であるという．正則でない行列は**特異**(singular)であるという．

2.10.2　逆行列，余因子行列に関する諸公式

注意 2.10.1 で指摘したように，逆行列は存在すれば一意であるから，逆行列に関する公式が成立することを示すには，単に“ある行列と他の行列との積が単位行列になる”ということを確認するだけでよい．以下その例を示す．

(ii)　$(AB)^{-1}=B^{-1}A^{-1}$　(2.10.4)

“A^{-1} も B^{-1} も存在するときには $(AB)^{-1}$ も存在して(2.10.4)が成立する”と読む(以下同様)．(2.10.4)が成り立つことは，$(AB)(B^{-1}A^{-1})=A(BB^{-1})A^{-1}=AIA^{-1}=AA^{-1}=I$ のようにして確認できる．

> **注意 2.10.2**　(2.2.18)，(2.2.19)で定義した正方行列 A のべきあるいは多項式は，A が正則なら，$A^{-p}=(A^{-1})^p$ とおくことによって，自然に負のべきにも拡張できる．

(iii)　$(A^{-1})^{-1}=A$　(2.10.5)

略証　$A^{-1}A=I$ の読み替えにすぎない．■

(iv)　$(A^{\mathsf{T}})^{-1}=(A^{-1})^{\mathsf{T}}$　　($K=\mathbb{C}$ のとき $(A^{\mathsf{H}})^{-1}=(A^{-1})^{\mathsf{H}}$)　(2.10.6)

略証　$AA^{-1}=I$ の両辺の転置を取れば．$(A^{-1})^{\mathsf{T}}A^{\mathsf{T}}=I$．■

(v)　A が n 次正方正則，D が m 次正方，B が $n\times m$ 型，C が $m\times n$ 型，$S\equiv D-CA^{-1}B$ が正則のとき，

$$\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right]^{-1}=\left[\begin{array}{c:c} A^{-1}+A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1} \\ \hdashline -S^{-1}CA^{-1} & S^{-1} \end{array}\right] \tag{2.10.7}$$

となる．特に $m=1$ のとき，$s=d-\boldsymbol{c}^{\mathsf{T}}A^{-1}\boldsymbol{b}\neq 0$ ならば，

$$\left[\begin{array}{c:c} A & \boldsymbol{b} \\ \hdashline \boldsymbol{c}^{\mathsf{T}} & d \end{array}\right]^{-1}=\left[\begin{array}{c:c} A^{-1}+(A^{-1}\boldsymbol{b})(\boldsymbol{c}^{\mathsf{T}}A^{-1})/s & -A^{-1}\boldsymbol{b}/s \\ \hdashline -\boldsymbol{c}^{\mathsf{T}}A^{-1}/s & 1/s \end{array}\right], \tag{2.10.8}$$

$$s=d-\boldsymbol{c}^{\mathsf{T}}A^{-1}\boldsymbol{b}.$$

略証　$F=\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right]$ および $W=\left[\begin{array}{c:c} I_n & O_{n,m} \\ \hdashline -CA^{-1} & I_m \end{array}\right]$ とおくと，$G=WF=$

$\left[\begin{array}{c:c} A & B \\ \hdashline O_{m,n} & S \end{array}\right]$. さらに，$U=\left[\begin{array}{c:c} I_n & -BS^{-1} \\ \hdashline O_{m,n} & I_m \end{array}\right]$とおくと，$H=UG=UWF=\left[\begin{array}{c:c} A & O_{n,m} \\ \hdashline O_{m,n} & S \end{array}\right]$. ゆえに，$H^{-1}=F^{-1}W^{-1}U^{-1}$，すなわち，$F^{-1}=H^{-1}UW$. $H^{-1}=\left[\begin{array}{c:c} A^{-1} & O_{n,m} \\ \hdashline O_{m,n} & S^{-1} \end{array}\right]$より(2.10.7)が導かれる. ■

注意 2.10.3　$S=D-CA^{-1}B$ のことを，$\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right]$の($A$ を軸とする)**Schur**(シュール)**の補行列**(Schur's complement)と呼ぶことがある.

(vi)　(v)と同じような条件のもとで，

$$\det\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right]=(\det A)\cdot\det(D-CA^{-1}B)$$
$$=(\det D)\cdot\det(A-BD^{-1}C)\,. \qquad (2.10.9)$$

略証　(v)の証明の中で，$\det W=1$ であることに注意すれば，$G=WF$ より $\det G=(\det A)\cdot(\det S)=(\det W)\cdot(\det F)=\det F$. ■

(vii)　A が n 次正方正則，D が m 次正方正則，B が $n\times m$ 型，C が $m\times n$ 型，$D^{-1}+CA^{-1}B$ が正則のとき，

$$(A+BDC)^{-1}=A^{-1}-A^{-1}B(D^{-1}+CA^{-1}B)^{-1}CA^{-1}$$
$$=A^{-1}-A^{-1}BD(D+DCA^{-1}BD)^{-1}DCA^{-1}. \qquad (2.10.10)$$

略証　(2.10.7)で A と D の役割を入れ換えると，

$$\left[\begin{array}{c:c} A & B \\ \hdashline C & D \end{array}\right]^{-1}=\left[\begin{array}{c:c} T^{-1} & -T^{-1}BD^{-1} \\ \hdashline -D^{-1}CT^{-1} & D^{-1}+D^{-1}CT^{-1}BD^{-1} \end{array}\right],\quad T\equiv A-BD^{-1}C$$

を得る. この式と(2.10.7)の左上ブロックを比較すると，

$$(A-BD^{-1}C)^{-1}=A^{-1}+A^{-1}B(D-CA^{-1}B)^{-1}CA^{-1}$$

となる. ここで，D の代りに $-D^{-1}$ を代入すれば(2.10.10)の第1式を得る. 第2式は，右辺の括弧の両側に $I_m=DD^{-1}, I_m=D^{-1}D$ を挿入し，(2.10.4)を用いれば得られる. ■

注意 2.10.4　(2.10.10)(の第1式)において $m=1$ とおき，$B=\boldsymbol{b}, C=\boldsymbol{c}^{\top}, D=d$ と記し，一般性を失うことなく $d=1$ とすると，

$$(A+\boldsymbol{bc}^{\mathsf T})^{-1}=A^{-1}-(A^{-1}\boldsymbol{b})(\boldsymbol{c}^{\mathsf T}A^{-1})/(1+\boldsymbol{c}^{\mathsf T}A^{-1}\boldsymbol{b}) \tag{2.10.11}$$

となる．この式は，正方正則行列 A を階数 1 の行列 $\boldsymbol{bc}^{\mathsf T}$(→2.13 節)で修正したものが正則であるための必要十分条件が $\boldsymbol{c}^{\mathsf T}A^{-1}\boldsymbol{b}\neq-1$ であることと，逆行列の変化の具体形を示している．より直接的に，(2.10.9)の中辺と右辺の関係を用いて，$B=-\boldsymbol{b}$ とおけば，

$$\det(A+\boldsymbol{bc}^{\mathsf T})=(\det A)\cdot(1+\boldsymbol{c}^{\mathsf T}A^{-1}\boldsymbol{b}) \tag{2.10.9$'$}$$

であることにも注意．なお，(2.10.11)を **Sherman-Morrison の公式**，(2.10.10)を **Sherman-Morrison-Woodbury の公式**と呼ぶ人もいる．

注意 2.10.5　行列や行列式に関する諸公式のうち，逆行列を含むもの(あるいは行列の正則性を前提とするもの)は特異行列に対しては成り立たないようにみえる．しかし，適当な行列式を両辺に掛けることによって(いわば分母を払って)得られる関係式はそのまま成立することが多い．厳密には次のように考えればよい．すなわち，A の n^2 個の要素 $a_{11},a_{12},\cdots,a_{nn}$ をみな独立な変数であるとみなせば(あるいは同じことだが，体 K 上で互いに代数的に独立な超越元であるとみなせば)，すべての式はそれらの変数の有理式全体からなる体 $K(a_{11},\cdots,a_{nn})$ (あるいは K に $a_{11},\cdots,a_{nn}$ を添加してできる超越拡大体)の上の式とみなせる．そこでは $\det A\neq0$ であるから，すべての正方行列は正則である．そこで，A の逆行列を含んだある関係式が $K(a_{11},\cdots,a_{nn})$ において成り立てば，$A^{-1}=\hat{A}/\det A$ とおいて，両辺に $\det A$ の適当なべきを掛けることによって，その関係式を $K(a_{11},\cdots,a_{nn})$ の“多項式の間の関係式”に書き直せる．そのように多項式の間の関係に書き直してしまえば，$a_{11},a_{12},\cdots,a_{nn}$ の中に K の元をどのように代入したものも正しい関係式となる．

上の注意 2.10.5 の考え方も援用しながら，小行列式，逆行列，余因子行列などの間に成り立つ有用な関係式の中で代表的なものをいくつか導いておこう(A は n 次正方行列)．

(viii)　$\det\hat{A}=(\det A)^{n-1}$ 　　(2.10.12)

略証　(2.10.3)に $A^{-1}=(\det A)^{-1}\hat{A}$ を代入して，(2.6.4)を用いれば，$(\det A)^{-n}\det\hat{A}=(\det A)^{-1}$．両辺に $(\det A)^n$ を掛ける．■

(ix)　$(\hat{A})^{-1}=\mathrm{adj}(A^{-1})=(\det A)^{-1}\cdot A$ 　　(2.10.13)

略証　$(\hat{A})^{-1}=(\det A)^{-1}\cdot A$ は(2.8.4)の書き換え．(2.10.5)，すなわち $A=(A^{-1})^{-1}=(\det A^{-1})^{-1}\cdot\mathrm{adj}(A^{-1})=(\det A)\cdot\mathrm{adj}(A^{-1})$ の両辺に $(\det A)^{-1}$ を掛ければ $\mathrm{adj}(A^{-1})=(\det A)^{-1}\cdot A$ を得る．■

(x)　$\mathrm{adj}(AB)=\widehat{B}\cdot\widehat{A}$　(2.10.14)

略証　(2.10.4)の両辺に $\det(AB)=(\det A)\cdot(\det B)$ を掛ける．■

(xi)　$\mathrm{adj}(A^{\mathsf{T}})=(\widehat{A})^{\mathsf{T}}$　($K=\mathbb{C}$ のとき　$\mathrm{adj}(A^{\mathsf{H}})=(\widehat{A})^{\mathsf{H}}$)　(2.10.15)

略証　(2.10.6)に $\det A=\det A^{\mathsf{T}}$ を掛ける．■

(xii)　$\mathrm{adj}\,\widehat{A}=(\det A)^{n-2}\cdot A$　(2.10.16)

略証　(2.10.12)，(2.10.13)より，$\mathrm{adj}\,\widehat{A}=(\det\widehat{A})\cdot(\widehat{A})^{-1}=(\det A)^{n-2}\cdot A$ を得る．■

(xiii)　$\mathrm{adj}(\lambda A)=\lambda^{n-1}\widehat{A}$　(2.10.17)

略証　余因子行列の各要素は A の $n-1$ 次の小行列式であるから．■

2.10.3 Jacobi 型の公式

$A,\ B=A^{-1}$ が n 次正方行列で$(r,n-r;r,n-r)$型に分割されているとする：

$$A=\begin{bmatrix}A_{11} & A_{12}\\ A_{21} & A_{22}\end{bmatrix},\quad B=\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}.\qquad(2.10.18)$$

(行ブロックの大きさ $r,\ n-r$；列ブロックの大きさ $r,\ n-r$)

$B=A^{-1}$ であるから，

$$\begin{bmatrix}A_{11} & A_{12}\\ O_{n-r,r} & I_{n-r}\end{bmatrix}\cdot\begin{bmatrix}B_{11} & B_{12}\\ B_{21} & B_{22}\end{bmatrix}=\begin{bmatrix}I_r & O_{r,n-r}\\ B_{21} & B_{22}\end{bmatrix}\qquad(2.10.19)$$

である．これより

$$(\det A_{11})\cdot(\det A^{-1})=\det B_{22},$$

すなわち

$$\det A_{11}=(\det A)\cdot(\det B_{22})\qquad(2.10.20)$$

が成り立つ．$\widehat{A}$ も同様の分割

$$\widehat{A}=\begin{bmatrix}\widehat{A}_{11} & \widehat{A}_{12}\\ \widehat{A}_{21} & \widehat{A}_{22}\end{bmatrix}\qquad(2.10.21)$$

で表すと，(2.10.20)は

$$(\det A)^{n-r-1}\cdot(\det A_{11})=\det\widehat{A}_{22}\qquad(2.10.22)$$

と書き直せる．行と列を適当に入れ換えれば，

(xiv)　$i_1<\cdots<i_r,\ \ i_{r+1}<\cdots<i_n,\ \ \{i_1,\cdots,i_r,i_{r+1},\cdots,i_n\}=\{1,\cdots,n\},$

$$j_1<\cdots<j_r,\ j_{r+1}<\cdots<j_n,\quad \{j_1,\cdots,j_r,j_{r+1},\cdots,j_n\}=\{1,\cdots,n\}$$

に対して，

$$|\hat{A}|^{i_{r+1}\cdots i_n}_{j_{r+1}\cdots j_n}=(\det A)^{n-r-1}|A|^{j_1\cdots j_r}_{i_1\cdots i_r}. \tag{2.10.23}$$

この形の関係式は **Jacobi（ヤコビ）の定理**と呼ばれることがある．特に，$r=n-2$，$i_k=j_k=k\,(k=1,\cdots,n)$の場合は，

$$\text{(xiv}'\text{)}\quad \det\begin{bmatrix}\varDelta_{n-1,n-1} & \varDelta_{n-1,n}\\ \varDelta_{n,n-1} & \varDelta_{n,n}\end{bmatrix}=|\hat{A}|^{n-1,n}_{n-1,n}=(\det A)\,|A|^{1,\cdots,n-2}_{1,\cdots,n-2} \tag{2.10.24}$$

という形になる．

2.10.4　Sylvester 型の公式

(2.10.9)（関係している行列の次数，型に注意）からは

$$\text{(xv)}\quad (\det A)^{m-1}\cdot\det\begin{bmatrix}\overset{\leftarrow n\rightarrow}{A} & \overset{\leftarrow m\rightarrow}{B}\\ C & D\end{bmatrix}\begin{matrix}\updownarrow n\\ \updownarrow m\end{matrix}=\det((\det A)D-C\hat{A}B) \tag{2.10.25}$$

が出る．d がスカラー（1×1 型行列），$\boldsymbol{b}$ が m 次元縦ベクトル（$m\times 1$ 型行列），$\boldsymbol{c}^{\mathsf{T}}$ が m 次元横ベクトル（$1\times m$ 型行列）のとき，(2.10.25)の特別な場合（$m=1$）として，**縁どり行列**(bordered matrix)の行列式の表現

$$\text{(xvi)}\quad \det\begin{bmatrix}A & \boldsymbol{b}\\ \boldsymbol{c}^{\mathsf{T}} & d\end{bmatrix}=(\det A)\cdot d-\boldsymbol{c}^{\mathsf{T}}\cdot\hat{A}\cdot\boldsymbol{b} \tag{2.10.26}$$

が得られる．(2.10.26)において，(2.10.25)の D の(i,j)要素を $d=d_{ij}$, B の列 j を $\boldsymbol{b}=\boldsymbol{b}_{\cdot j}$，$C$ の行 i を $\boldsymbol{c}^{\mathsf{T}}=\boldsymbol{c}_{i\cdot}$ $(i,j=n+1,\cdots,n+m)$とおくと，

$$\det\begin{bmatrix}A & \boldsymbol{b}_{\cdot j}\\ \boldsymbol{c}_{i\cdot} & d_{ij}\end{bmatrix}=(\det A)\cdot d_{ij}-\boldsymbol{c}_{i\cdot}\cdot\hat{A}\cdot\boldsymbol{b}_{\cdot j} \tag{2.10.27}$$

であるが，(2.10.27)の右辺は(2.10.25)の右辺の m 次正方行列$(\det A)\cdot D-C\hat{A}B$ の$(i-n,j-n)$要素に他ならないから，

$$\tilde{d}_{ij}=\det\begin{bmatrix}A & \boldsymbol{b}_{\cdot,n+j}\\ \boldsymbol{c}_{n+i,\cdot} & d_{n+i,n+j}\end{bmatrix}\qquad (i,j=1,\cdots,m) \tag{2.10.28}$$

を要素とする m 次正方行列 $\tilde{D}=[\tilde{d}_{ij}]$の行列式に対して，

$$\text{(xvii)}\quad \det\tilde{D}=(\det A)^{m-1}\cdot\det\begin{bmatrix}A & B\\ C & D\end{bmatrix} \tag{2.10.29}$$

という公式が得られる．この形の公式を **Sylvester の定理**ということがある．

(H. G. Grassmann によるともいわれる.)

2.10.5 行列式のもう一つの特徴づけ

n 次正方行列 A の行列式を A の n^2 個の要素 a_{ij} の関数(結局は多項式になるわけであるが)と見たときの特徴づけにはいくつかの種類がある. (2.6.1), (2.6.2)のような陽な定義の他に, 2.6 節の性質(i)〜(iv)による特徴づけも可能であることは既に見た(注意 2.6.3).

もう一つの特徴づけとして性質(2.8.16)によるものがある. すなわち, n 次正方行列 A の要素の多項式 $\varphi(A)$ が

$$\varphi(AB)=\varphi(A)\,\varphi(B) \tag{2.10.30}$$

を満たすとき, $\varphi(A)=(\det A)^k$ (k は自然数)である. φ が(2.10.30)を満たす定数でない最低次の多項式ならば, $\varphi(A)=\det A$ である. (φ が有理式, 連続関数, などなどの場合への一般化も考えられるが, ここでは簡単のため多項式ということにしておく.)

実際, $A=\lambda I_n$ ($\lambda\in K$) として, $p(\lambda)=\varphi(\lambda I_n)$ (λ の多項式)とおくと, $p(\lambda\mu)=\varphi(\lambda\mu I_n)=\varphi((\lambda I_n)\cdot(\mu I_n))=\varphi(\lambda I_n)\,\varphi(\mu I_n)=p(\lambda)\,p(\mu)$. $p(\lambda\mu)=p(\lambda)\,p(\mu)$ を満たす多項式 p は $p(\lambda)=\lambda^h$ (h は自然数)に限られる. ($p(\lambda)=\sum_{i=0}^{h}c_i\lambda^i$ ($c_h\neq 0$) とすると, $p(\lambda\mu)=\sum_{i=0}^{h}c_i\lambda^i\mu^i=p(\lambda)\,p(\mu)=\sum_{i=0}^{h}\sum_{j=0}^{h}c_ic_j\lambda^i\mu^j$ となるが, 両辺の μ^h の係数を比較して, $c_h\lambda^h=\left(\sum_{i=0}^{h}c_i\lambda^i\right)c_h$, すなわち $c_h=1$, $c_i=0$ ($i\neq h$), ゆえに $p(\lambda)=\lambda^h$ でなければならない.)

次に, (2.8.4)を用いれば $\varphi(A)\,\varphi(\hat{A})=\varphi((\det A)\,I_n)=p(\det A)=(\det A)^h$ となる. $\hat{A}$ の要素が a_{ij} の多項式であるから $\varphi(\hat{A})$ も結局 a_{ij} の多項式である. 一方, 2.6.2 項により, $\det A$ は a_{ij} の既約多項式であるから, $k+l=h$ であるようなある k,l に対して $\varphi(A)=c\,(\det A)^k$, $\varphi(\hat{A})=c^{-1}(\det A)^l$ となっていなければならない. 再び $A=\lambda I_n$ とすれば, $\varphi(\lambda I_n)=c\lambda^{nk}=p(\lambda)=\lambda^h$. ゆえに, $c=1$, $h=nk$. したがって, $\varphi(A)=(\det A)^k$ でなければならない.

逆に, $\varphi(A)=(\det A)^k$ (k は自然数)が(2.10.30)を満たすことは明らかである.

2.10.6 Kronecker 積の行列式

(xviii)　m 次正方行列 A と n 次正方行列 B との Kronecker 積 $A\otimes B$

(→2.4.2 項)の行列式は，

$$\det(A\otimes B)=(\det A)^n\cdot(\det B)^m \tag{2.10.31}$$

である．

略証　$A\otimes B=(I_m\cdot A)\otimes(B\cdot I_n)=(I_m\otimes B)\cdot(A\otimes I_n)$ と書き表しておく(→(2.4.6))．$I_m\otimes B=\left.\begin{bmatrix} B & & O \\ & \ddots & \\ O & & B \end{bmatrix}\right\} m$ 回　であるから，$\det(I_m\otimes B)=(\det B)^m$．$A\otimes I_n$ は $I_n\otimes A$ の行と列に同じ順序の入れ換えをおこなったものだから，

$$\det(A\otimes I_n)=\det(I_n\otimes A)=(\det A)^n$$

となる．■

2.11　同次の小行列式の間の関係

$m\times n$ 型の行列の要素の総数が mn 個であるのに対して，その r 次の小行列式は(符号は別として自明に一致するものを除いて)$\binom{m}{r}\cdot\binom{n}{r}$個ある．したがって，小行列式の間に成り立つ多くの関係があるはずである．

2.11.1　Grassmann-Plücker の恒等式

$m\times n$ 型の行列 $A(m\leqq n)$ の m 次の小行列式 $|A|^{1\cdots m}_{j_1\cdots j_m}$ $(j_k=1,\cdots,n)$ の列番号のみに注目して，

$$v_{j_1\cdots j_m}=|A|^{1\cdots m}_{j_1\cdots j_m} \tag{2.11.1}$$

とおくと，任意の $i_1,\cdots,i_m$，$j_1,\cdots,j_m$ $(=1,\cdots,n)$ に対して

$$v_{[j_1\cdots j_m}v_{i_1]\,i_2\cdots i_m}=0 \tag{2.11.2}$$

が成り立つ．(2.11.2)は **Grassmann-Plücker**(グラスマン-プリュッカー)**の恒等式**と呼ばれている．(2.11.1)の $v_{j_1\cdots j_m}$ は，添字 $j_1,\cdots,j_m$ に関して交代的である(行列式の定義(2.6.2)の形より明らか)：

$$v_{j_1\cdots j_m}=v_{[j_1\cdots j_m]}. \tag{2.11.3}$$

(2.5.10)によれば，(2.11.2)は($m+1$ 倍して)

$$v_{j_1\cdots j_m}v_{i_1 i_2\cdots i_m}=\sum_{k=1}^{m} v_{j_1\cdots j_{k-1}i_1 j_{k+1}\cdots j_m}v_{j_k i_2\cdots i_m} \tag{2.11.4}$$

とも表せる．

(2.11.2)が成り立つことは，行列式の定義(2.6.2)と交代化演算の性質(2.5.5)によりほとんど自明である．すなわち，

$$v_{[j_1\cdots j_m}v_{i_1]i_2\cdots i_m}=(m!)^2a^{[1}_{[j_1}\cdots a^{m]}_{j_m}a^{[1}_{i_1]}a^{2}_{i_2}\cdots a^{m]}_{i_m}$$
$$=(m!)^2a^{[\![[1}_{[j_1}\cdots a^{m]}_{j_m}a^{[1]\!]}_{i_1]}a^{2}_{i_2}\cdots a^{m]}_{i_m} \qquad (2.11.5)$$

であるが，$[1\cdots m][1\cdots m]$の右の方の$[1\cdots m]$のどの置換(たとえば$k(=1,\cdots,m)$が先頭に来るもの)に対しても，交代化$[\![1\cdots m\ k]\!]$において，$1\cdots m$のどれかとkが等しいので，最右辺の和は0となってしまうからである．

注意 2.11.1　いうまでもないことであるが，Grassmann-Plückerの恒等式は，行列Aの行と列の役割を入れ換えたものについても成立するし，また，行列Aの行(あるいは列)を共有する部分行列の小行列式に対しても成立する．

注意 2.11.2　(2.11.2)あるいは同じことであるが，(2.11.4)は，このように，見方によっては自明な性質であるが，歴史的には，射影幾何学，微分幾何学において重要な役割を果たした関係式である．また，この恒等式から導き出される諸性質をもとにして**マトロイド**(matroid)あるいは**有向マトロイド**(oriented matroid)と呼ばれる数学的体系を構築することによって，応用上も有用な組合せ数学の一大分野が展開されるという意味でも重要な性質である．

注意 2.11.3　交代的な量$v_{j_1\cdots j_m}=v_{[j_1\cdots j_m]}$ $(j_k=1,2,\cdots,n)$が(2.11.2)を満たし，かつすべては0でないとする．たとえば，一般性を失わずに$v_{1\cdots m}\neq 0$であるとする．すると，$v_{1\cdots m}$と添字が1カ所だけ異なるような

$$v_{j2\cdots m},\ v_{1j3\cdots m},\ \cdots,\ v_{12\cdots,k-1,j,k+1,\cdots,m},\ \cdots,\ v_{12\cdots,m-1,j} \qquad (j=m+1,\cdots,n) \qquad (2.11.6)$$

と$v_{1\cdots m}$とによって，他のすべての$v_{j_1\cdots j_m}$が一意に定まる．実際，$v_{j_1\cdots j_m}$に対して$\{j_1,\cdots,j_m\}$と$\{1,\cdots,m\}$が共通の要素を$m-2$個以下しか有さないとき，すなわち，

$$|\{j_1,\cdots,j_m\}\cap\{1,\cdots,m\}|\leqq m-2 \qquad (2.11.7)$$

のとき，任意の$j\in\{j_1,\cdots,j_m\}-\{1,\cdots,m\}$(一般性を失わずに，$j=j_1$であるとする)に対して，(2.11.4)より

$$v_{1\cdots m}v_{j_1j_2\cdots j_m}=\sum_{k=1}^{m}v_{1,\cdots,k-1,j_1,k+1,\cdots,m}v_{kj_2\cdots j_m} \qquad (2.11.8)$$

であるが，(2.11.8)の右辺の$\sum$は$k\in\{1,\cdots,m\}-\{j_2,\cdots,j_m\}$についてのみ取ればよい．(それ以外の$k$に対しては，$v_{kj_2\cdots j_m}=0$である！)右辺の$v_{1,\cdots,k-1,j_1,k+1,\cdots,m}$は(2.11.6)のどれかである．一方，$j_1\in\{j_1,\cdots,j_m\}-\{1,\cdots,m\}$，$k\in\{1,\cdots,m\}-\{j_2,\cdots,j_m\}$であるから，右辺の$v_{kj_2\cdots j_m}$と$v_{1\cdots m}$とに共通の添字の個数は，明らかに，もとの$v_{j_1\cdots j_m}$と$v_{1\cdots m}$とに共通の添字の個数より1個だけ多い．(2.11.8)の両辺を$v_{1\cdots m}(\neq 0)$で割れば，$v_{j_1\cdots j_m}$を$v_{1\cdots m}$と(2.11.6)の形のvと

$$|\{j_1',\cdots,j_m'\}\cap\{1,\cdots,m\}|=|\{j_1,\cdots,j_m\}\cap\{1,\cdots,m\}|+1$$

であるような$v_{j_1'\cdots j_m'}$とで表す式となる．これを繰り返せば，$v_{j_1\cdots j_m}$を$v_{1\cdots m}$と

▒　(2.11.6)の形の v とで表す式が得られる．

$m\times n$ 型の行列 $A(m\leqq n)$ の小行列式(2.11.1)が(2.11.2)を満たすことは，このように容易に示せるが，そのある意味での逆も成り立つ．すなわち，交代的な量 $v_{j_1\cdots j_m}=v_{[j_1\cdots j_m]}$ $(j_k=1,\cdots,n)$ が(2.11.2)を満たすならば，ある $m\times n$ 型の行列 A が存在して，$v_{j_1\cdots j_m}$ は A の小行列式(2.11.1)に等しい．このことは以下のように示せる．

すべての $v_{j_1\cdots j_m}=0$ であるなら $A=O_{m,n}$ とすればよい．そうでないときには，一般性を失うことなく，$v_{1\cdots m}\neq 0$ であるとしてよい．やや天下り的であるが，

$$
\begin{aligned}
&a^1_j=v_{j2\cdots m} \qquad (j=1,\cdots,n),\\
&a^i_j=v_{1,\cdots,i-1,j,i+1,\cdots,m}/v_{1\cdots m} \qquad (i=2,\cdots,m\,;\,j=1,\cdots,n)
\end{aligned}
\tag{2.11.9}
$$

とおく．明らかに

$$
\begin{aligned}
&a^1_1=v_{1\cdots m},\quad a^1_j=0 \qquad (j=2,\cdots,m),\\
&a^i_i=1,\quad a^i_j=0 \qquad (i=2,\cdots,m\,;\,j\neq i,j=1,\cdots,m)
\end{aligned}
\tag{2.11.10}
$$

である．ゆえに，

$$
\begin{aligned}
|A|^{12\cdots m}_{j2\cdots m}&=v_{j2\cdots m} \qquad (j=1,\cdots,n),\\
|A|^{1,\cdots,i-1,i,i+1,\cdots,m}_{1,\cdots,i-1,j,i+1,\cdots,m}&=v_{1\cdots m}\cdot v_{1,\cdots,i-1,j,i+1,\cdots,m}/v_{1\cdots m}\\
&=v_{1,\cdots,i-1,j,i+1,\cdots,m}\\
&\qquad (i=2,\cdots,m\,;\,j=1,\cdots,n)
\end{aligned}
\tag{2.11.11}
$$

である．このように，$\{1,\cdots,m\}$ と要素がたかだか1個しか異ならないようなすべての $\{j_1,\cdots,j_m\}$ $(j_k=1,\cdots,n)$ に対して(2.11.1)が成立しているので，上の注意2.11.3により，すべての $\{j_1,\cdots,j_n\}$ に対して(2.11.1)が成り立つ．

2.11.2　Kronecker-Runge の恒等式

Grassmann-Plücker の恒等式は一般の行列の小行列式の間の“2次の”関係式であったが，n 次**対称**(正方)行列の r 次の小行列式の間には次のような“1次の”関係式(Kronecker の恒等式，あるいは **Kronecker-Runge**(クロネッカー-ルンゲ)**の恒等式**)が成立する．

$$
|A|^{i_1\cdots i_{r-1}i_r}_{j_1\cdots\cdots j_r}=\sum_{k=1}^{r}|A|^{i_1\cdots\cdots\cdots i_{r-1}j_k}_{j_1\cdots j_{k-1}i_rj_{k+1}\cdots j_r} \qquad (i_1,\cdots,i_r,j_1,\cdots,j_r=1,2,\cdots,n).
\tag{2.11.12}
$$

(2.11.12)が成立することは容易に示せる．実際，$|A|^{i_1\cdots i_r}_{j_1\cdots j_r}$ を交代化演算を用いて

$$|A|^{i_1\cdots i_r}_{j_1\cdots j_r}=r!\,a_{i_1[j_1}a_{|i_2|j_2}\cdots a_{|i_{r-1}|j_{r-1}}a_{|i_r|j_r]} \tag{2.11.13}$$

と表しておくと，(2.11.12)の左辺から右辺を引いたものは，(2.5.10)により，

$$(r+1)!\,a_{i_1[j_1}a_{|i_2|j_2}\cdots a_{|i_{r-1}|j_{r-1}}a_{i_rj_r]} \tag{2.11.14}$$

と書ける((2.11.13)では i_r が交代化の範囲に入っていないが，(2.11.14)では入っていることに注意)．(2.11.14)において，$j_1, j_2, \cdots, j_r, i_r$ のすべての置換を書き下すと，最後の $a_{i_rj_r}$ のところが a_{ij} $(i, j=j_1, j_2, \cdots, j_r, i_r)$ である項と a_{ji} である項(そしてそれ以外の因子は同じ)とが反対符号で現れる．A が対称，すなわち，$a_{[ij]}=0$ であるから，(2.11.14)は恒等的に 0 に等しい．

このように(2.11.12)が成り立つことは明らかであるが，実は，

「対称行列の小行列式の間に成り立つ(自明でない)すべての 1 次の関係式は，(2.11.12)の形の関係式を足したり引いたりして得られる」

ということも知られている．(証明は初等的で困難ではないが，長くなるのでここでは省く．)

注意 2.11.4 反対称(正方)行列に対しては，対称行列に対する Kronecker-Runge の恒等式に対応する 1 次の関係式は**存在しない**ことが知られている．

注意 2.11.5 ほとんど自明なことであるが，一般の $m\times n$ 型行列 A に対して，その r 次の小行列式($A^{<r>}$ の要素)の間に成り立つ“1 次の”関係式は存在しない．なぜならば，もしある係数 $c^{i_1\cdots i_r}_{j_1\cdots j_r}$ を用いて

$$\sum c^{i_1\cdots i_r}_{j_1\cdots j_r}|A|^{i_1\cdots i_r}_{j_1\cdots j_r}=0$$
$$(i_1<\cdots<i_r\,;\,j_1<\cdots<j_r\,;\,i_k=1,\cdots,m\,;\,j_k=1,\cdots,n) \tag{2.11.15}$$

という形の関係が任意の行列 A に対して成立したとすると，それは特に，

$$a^1_1=a^2_2=\cdots=a^r_r=1,\quad \text{その他の } a^i_j=0 \tag{2.11.16}$$

という行列 A に対しても成立しなければならないが，このような A の r 次の小行列式の中で 0 にならないものは，$|A|^{1\cdots r}_{1\cdots r}$ $(=1)$だけであるので，(2.11.15)において $c^{1\cdots r}_{1\cdots r}=0$ でなければならないことになる．一般の $m\times n$ 型行列については行番号と列番号は独立に振り換えて差し支えないから，他の $c^{i_1\cdots i_r}_{j_1\cdots j_r}$ も同様にしてすべて 0 でなければならない．

2.12 極限と微積分

2.12.1 行列級数

$K=\mathbb{R}$，$K=\mathbb{C}$ などの場合，$m\times n$ 型の行列の系列 $A^{(1)}=[a_{ij}^{(1)}]$，$A^{(2)}=[a_{ij}^{(2)}]$，$\cdots$ の極限 $A^{(\infty)}=[a_{ij}^{(\infty)}]$への収束は，各要素ごとの収束で定義される．

$$\lim_{k\to\infty} A^{(k)}=A^{(\infty)} \Longleftrightarrow \lim_{k\to\infty} a_{ij}^{(k)}=a_{ij}^{(\infty)}$$
$$(i=1,\cdots,m\,;\, j=1,\cdots,n)\,. \qquad (2.12.1)$$

微積分学の初歩の教えるところによれば，下記の演算が意味をもつような型の行列に対して，

$$\lim_{k\to\infty}(\alpha A^{(k)}+\beta B^{(k)})=\alpha\lim_{k\to\infty} A^{(k)}+\beta\lim_{k\to\infty} B^{(k)} \qquad (\alpha,\beta\in K), \quad (2.12.2)$$

$$\lim_{k\to\infty}(A^{(k)}\cdot B^{(k)})=\left(\lim_{k\to\infty} A^{(k)}\right)\cdot\left(\lim_{k\to\infty} B^{(k)}\right) \qquad (2.12.3)$$

が成り立つ．

n 次正方行列 A に対して，形式和

$$B=a_0I+a_1A+a_2A^2+\cdots=\sum_{k=0}^{\infty}a_kA^k \qquad (2.12.4)$$

を A の**べき級数**(power series)と呼ぶ($a_0,a_1,\cdots\in K$)．部分和である行列多項式

$$B^{(k)}=a_0I+a_1A+a_2A^2+\cdots+a_kA^k \qquad (2.12.5)$$

が極限をもつ，すなわち $\lim_{k\to\infty} B^{(k)}=B^{(\infty)}$ のとき，$B=B^{(\infty)}$ であると定義し，べき級数(2.12.4)は収束するという．絶対収束などについても，要素ごとのそれとして同様に定義される．

行列 A のべき級数の収束性は，対応する 1 変数 x のべき級数

$$f(x)=a_0+a_1x+a_2x^2+\cdots=\sum_{k=0}^{\infty}a_kx^k \qquad (2.12.6)$$

の収束半径と A の固有値(→第 5 章)との関係で定まるが，Jordan 標準形(→6.8 節)を仲介として理解するのが早道であるので，ここでは詳細には立ち入らない．

2.12.2 微　分

行列 $A=[a_{ij}]$の要素 a_{ij} がある変数 t の関数 $a_{ij}=a_{ij}(t)$であるとき，$A=A(t)$と記す．(2 個以上の変数の関数である場合も同様であるが，簡単のため，

ここでは1変数の場合について述べる.)

すべての要素が t に関して微分可能であるとき，$a_{ij}(t)$ の導関数を要素とする行列を $\mathrm{d}A(t)/\mathrm{d}t$ あるいは $A'(t)$ と書く：

$$A'(t)=\frac{\mathrm{d}}{\mathrm{d}t}A(t)=\left[\frac{\mathrm{d}}{\mathrm{d}t}a_{ij}(t)\right]. \tag{2.12.7}$$

演算が意味をもつような型の行列に対して，

$$\frac{\mathrm{d}}{\mathrm{d}t}(\alpha A(t)+\beta B(t))=\alpha\frac{\mathrm{d}}{\mathrm{d}t}A(t)+\beta\frac{\mathrm{d}}{\mathrm{d}t}B(t) \quad (\alpha,\beta\in K), \tag{2.12.8}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(A(t)\cdot B(t))=\left(\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right)\cdot B(t)+A(t)\cdot\left(\frac{\mathrm{d}}{\mathrm{d}t}B(t)\right) \tag{2.12.9}$$

であることは明らかであろう.

n 次正方行列 A の行列式 $\det A$ を A の n^2 個の要素を変数とする関数とみなすと，Laplace 展開の公式(2.8.1)あるいは(2.8.2)により，

$$\frac{\partial}{\partial a_{ij}}\det A=\Delta_{ij}=\hat{a}_{ji} \tag{2.12.10}$$

であることが直ちにわかる(a_{ij} の余因子 $\Delta_{ij}=\hat{a}_{ji}$ には a_{ij} が含まれていないから). (2.12.10)は，両辺を $\det A$ で割って

$$\frac{\partial}{\partial a_{ij}}(\log(\det A))=b_{ji} \qquad (B=[b_{ij}]=A^{-1}) \tag{2.12.11}$$

という形にして使った方が便利なことも多い.

n 次正方行列 $A(t)$ の逆行列 $A(t)^{-1}$ の導関数については，

$$A(t)A(t)^{-1}=I \tag{2.12.12}$$

(右辺は定数であることに注意！)に対して，(2.12.9)を適用して得られる式

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right)\cdot A(t)^{-1}+A(t)\cdot\left(\frac{\mathrm{d}}{\mathrm{d}t}A(t)^{-1}\right)=O \tag{2.12.13}$$

の両辺に左から $A(t)^{-1}$ を掛けると，

$$\frac{\mathrm{d}}{\mathrm{d}t}A^{-1}(t)=-A(t)^{-1}\cdot\left(\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right)\cdot A(t)^{-1} \tag{2.12.14}$$

という表式が得られる. あるいは，微分(differential)を用いた形で

$$\mathrm{d}A^{-1}=-A^{-1}\cdot\mathrm{d}A\cdot A^{-1} \tag{2.12.15}$$

と表すと便利なことがある.

特に，$B=[b_{ij}]=A^{-1}$ の要素 b_{ij} を $A=[a_{ij}]$ の要素 a_{ij} を変数とする関数とみたとき，(2.12.15)は

$$\frac{\partial b_{ij}}{\partial a_{kl}} = -b_{ik}b_{lj} \tag{2.12.16}$$

を表している.

n 次正方行列 $A(t)$ の行列式 $\det A(t)$ については,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\det A(t)) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial}{\partial a_{ij}}(\det A(t))\cdot\frac{\mathrm{d}}{\mathrm{d}t}a_{ij}(t) \tag{2.12.17}$$

であるから,(2.12.10)と Laplace 展開の公式(2.8.1)あるいは(2.8.2)を用いて,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}(\det A(t)) &= \sum_{i=1}^{n}\sum_{j=1}^{n}\Delta_{ij}(t)\cdot\frac{\mathrm{d}}{\mathrm{d}t}a_{ij}(t)\\ &= \sum_{j=1}^{n}\det\left[\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot j-1},\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{a}_{\cdot j},\boldsymbol{a}_{\cdot j+1},\cdots,\boldsymbol{a}_{\cdot n}\right]\\ &= \sum_{i=1}^{n}\det\begin{bmatrix}\boldsymbol{a}_{1\cdot}\\ \vdots\\ \boldsymbol{a}_{i-1\cdot}\\ \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{a}_{i\cdot}\\ \boldsymbol{a}_{i+1\cdot}\\ \vdots\\ \boldsymbol{a}_{n\cdot}\end{bmatrix}\end{aligned} \tag{2.12.18}$$

を得る.(2.12.18)は,要するに,「$\det A(t)$ の微分は,$A(t)$ の各列(あるいは各行)をその導関数でおきかえてできる行列の行列式の和である」ということを表している.

(2.12.18)の第1の等式は,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\det A(t)) = \mathrm{tr}\left(\hat{A}(t)\cdot\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right) \tag{2.12.19}$$

あるいは,$\det A(t) \neq 0$ のときには,

$$\frac{\mathrm{d}}{\mathrm{d}t}(\det A(t)) = (\det A(t))\cdot\mathrm{tr}\left(A(t)^{-1}\cdot\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right), \tag{2.12.20}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}(\log(\det A(t))) = \mathrm{tr}\left(A(t)^{-1}\cdot\frac{\mathrm{d}}{\mathrm{d}t}A(t)\right) \tag{2.12.21}$$

とも書き表せる.

例 2.12.1 n 次正方行列

$$A = I - B \tag{2.12.22}$$

に対して，行列級数

$$I+B+B^2+\cdots \tag{2.12.23}$$

が**絶対収束するならば**，それは A^{-1} に等しい．この級数は **Neumann**（ノイマン）**級数**と呼ばれる． □

例 2.12.2 n 次正方行列 A に対して，級数

$$I+\frac{1}{1!}A+\frac{1}{2!}A^2+\cdots=\sum_{k=0}^{\infty}\frac{1}{k!}A^k \tag{2.12.24}$$

は必ず**絶対収束する**．実際，$A=[a_{ij}]$ の絶対値最大の要素の絶対値を μ とすれば（$|a_{ij}|\leqq\mu$），A^k の要素の絶対値は $n^{k-1}\mu^k(\leqq(n\mu)^k)$ を超えないから，級数は絶対収束する．(2.12.24)の級数で定義される行列を $\exp(A)$ と記す．行列 A,B が**交換可能**，すなわち $AB=BA$ ならば，

$$\exp(A+B)=(\exp(A))\cdot(\exp(B)) \tag{2.12.25}$$

であることも容易に示せる．（これが記号 $\exp(A)$ を用いる理由である．）特に，(2.12.24)の表現からほとんど機械的に示せるように，

$$\exp(O_{n,n})=I_n, \qquad \exp((t+s)A)=(\exp(tA))\cdot(\exp(sA)),$$
$$(\exp(A))^{-1}=\exp(-A), \tag{2.12.26}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA)=A\cdot(\exp(tA))=(\exp(tA))\cdot A, \tag{2.12.27}$$

$$(\exp(A))^{\mathsf{T}}=\exp(A^{\mathsf{T}}), \qquad (\exp(A))^{\mathsf{H}}=\exp(A^{\mathsf{H}}) \tag{2.12.28}$$

なども成り立つ．また，(2.12.20)，(2.12.27)により，

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\det(\exp(tA))&=(\det(\exp(tA)))\cdot\mathrm{tr}((\exp(-tA))\cdot(\exp(tA))\cdot A)\\&=(\det(\exp(tA)))\cdot\mathrm{tr}(I\cdot A)\\&=(\det(\exp(tA)))\cdot(\mathrm{tr}\,A)\end{aligned} \tag{2.12.29}$$

であるが，(2.12.29)を $f(t)=\det(\exp(tA))$ に関する微分方程式とみなして $f(0)=\det I=1$ であること（(2.12.26)より）に注意すれば，

$$\det(\exp(tA))=\exp(t\cdot(\mathrm{tr}\,A)) \tag{2.12.30}$$

であることがわかる． □

2.13　階数，項別階数

2.13.1　階　　数

$m\times n$ 型の行列 A の r 次の小行列式は，それらすべてが 0 に等しいか，あるいはそれらの中に 0 でないものがあるかのどちらかである．0 でない小行列式があるような次数 r の最大値を A の**階数**(rank)といい，rank A で表す．

もちろん，$0\leqq \mathrm{rank}\,A\leqq \min(m,n)$ である．また，$\mathrm{rank}\,A=0$ は $A=O$ を意味する．さらに，A が n 次正方行列のとき，A が正則であることと $\mathrm{rank}\,A=n$ であることとは同じことである．

注意 2.13.1　行列の階数は，線形代数において最も重要な概念であって，ここでの定義の他に，多くの等価な特徴づけがある．たとえば，(1) 基本変形に基づく具体的計算によるもの(→ 2.15 節)，(2) 列(行)ベクトルの一次独立性によるもの(→ 3.1 節)，(3) 線形写像の像空間の次元によるもの(→ 3.3 節)，などである．特定の行列(写像)の階数を決定するには，場合に応じて適切な定義を用いるのがよいが，数値で要素が与えられた行列に対しては，一般には，(1)によるのがよい．m や n が大きいとき，上の定義は，階数を具体的に定める算法としてはほとんど実用にならない．

階数の定義より直ちにわかるように，

(i)　$\mathrm{rank}\,A^{\mathsf{T}}=\mathrm{rank}\,A$,

$K=\mathbb{C}$ のとき　$\mathrm{rank}\,\bar{A}=\mathrm{rank}\,A$,

$K=\mathbb{C}$ のとき　$\mathrm{rank}\,A^{\mathsf{H}}=\mathrm{rank}\,A$.　　(2.13.1)

A の r 次の小行列式がすべて 0 であるならば $r+1$ 次の小行列式もすべて 0 になることは，Laplace 展開(→ 2.8.1 項)によれば明らかである．そこで，

(ii)　$r\leqq \mathrm{rank}\,A$ に対しては，$A^{<r>}\neq O$.

$r>\mathrm{rank}\,A$ に対しては，$A^{<r>}=O$.

(ここで，$A^{<r>}$ は，A の r 次の小行列式を要素とする $\binom{m}{r}\times\binom{n}{r}$ 型の行列(→ 2.7.1 項).)

$C=AB$ のとき，C の一つの r 次の小行列式は，Binet-Cauchy 展開により，A の r 次の小行列式と B の r 次の小行列式との積和の形に表せるから(→

2.8.3 項)，$r>\min(\mathrm{rank}\,A, \mathrm{rank}\,B)$ ならば，$C^{<r>}=O$ である．すなわち，

(iii)　$\mathrm{rank}(AB) \leqq \min(\mathrm{rank}\,A, \mathrm{rank}\,B)$.　(2.13.2)

A が $m\times n$ 型の行列で，S が m 次正方正則行列($\mathrm{rank}\,S=m$)のときには，$\mathrm{rank}(SA) \leqq \min(m, \mathrm{rank}\,A)=\mathrm{rank}\,A$ であるが，上の(iii)を $A=S^{-1}\cdot(SA)$ に適用すると，$\mathrm{rank}\,A \leqq \min(m, \mathrm{rank}(SA))=\mathrm{rank}(SA)$ でもある．ゆえに，$\mathrm{rank}\,A=\mathrm{rank}(SA)$．$n$ 次正方正則行列 T についての $\mathrm{rank}(AT)$ に関しても同様．そこで，

(iv)　S, T が正則なら，

$$\mathrm{rank}\,A=\mathrm{rank}(SA)=\mathrm{rank}(AT)=\mathrm{rank}(SAT) . \tag{2.13.3}$$

A, B が同じ $m\times n$ 型の行列で，$\mathrm{rank}\,A=r$，$\mathrm{rank}\,B=s$，$r+s<\min(m, n)$ であるとき，$A+B$ の $t=r+s+1$ 次のある正方部分行列を選ぶと，その行列式は $\det[\tilde{\boldsymbol{a}}_{\cdot 1}+\tilde{\boldsymbol{b}}_{\cdot 1}, \cdots, \tilde{\boldsymbol{a}}_{\cdot t}+\tilde{\boldsymbol{b}}_{\cdot t}]$ と書ける($\tilde{\boldsymbol{a}}_{\cdot i}$, $\tilde{\boldsymbol{b}}_{\cdot i}$ は，それぞれ A あるいは B の，選ばれた行の要素からなる選ばれた列の列ベクトル)．行列式の性質(2.6.5)により，これは

$$\det[\tilde{\boldsymbol{a}}_{\cdot 1}+\tilde{\boldsymbol{b}}_{\cdot 1}, \cdots, \tilde{\boldsymbol{a}}_{\cdot t}+\tilde{\boldsymbol{b}}_{\cdot t}]=\sum_{\tilde{\boldsymbol{c}}_{\cdot i}=\tilde{\boldsymbol{a}}_{\cdot i}\ \text{あるいは}\ \tilde{\boldsymbol{b}}_{\cdot i}} \det[\tilde{\boldsymbol{c}}_{\cdot 1}, \cdots, \tilde{\boldsymbol{c}}_{\cdot t}] \tag{2.13.4}$$

と書き表せる．(2.13.4)の右辺の和の各項の $\tilde{\boldsymbol{c}}_{\cdot 1}, \cdots, \tilde{\boldsymbol{c}}_{\cdot t}$ の中には，$r+1$ 個以上の $\tilde{\boldsymbol{a}}_{\cdot i}$ が含まれるか，あるいは，$s+1$ 個以上の $\tilde{\boldsymbol{b}}_{\cdot i}$ が含まれるかのどちらかである．そこで，各項の行列式を，$\tilde{\boldsymbol{a}}_{\cdot i}$ の列と $\tilde{\boldsymbol{b}}_{\cdot i}$ の列に分けて一般化 Laplace 展開(2.8.6)を適用すれば(ただし，行と列の役割を入れ換える)，いずれの場合も，$\det[\tilde{\boldsymbol{c}}_{\cdot i}]=0$ となることがわかる．そこで次の関係が成り立つことがわかる(これは，$\mathrm{rank}\,A+\mathrm{rank}\,B \geqq \min(m, n)$ のときには自明に成立)．

(v)　$\mathrm{rank}(A+B) \leqq \mathrm{rank}\,A+\mathrm{rank}\,B$.　(2.13.5)

行列の積の階数が満たす重要な関係式には，以上の他に，

(vi)　A が $k\times l$ 型，B が $l\times m$ 型，C が $m\times n$ 型であるとき，

$$\mathrm{rank}(AB)+\mathrm{rank}(BC) \leqq \mathrm{rank}\,B+\mathrm{rank}(ABC) , \tag{2.13.6a}$$

$$\mathrm{rank}(BC)-\mathrm{rank}(ABC) \leqq \mathrm{rank}\,B-\mathrm{rank}(AB) , \tag{2.13.6b}$$

$$\mathrm{rank}(AB)-\mathrm{rank}(ABC) \leqq \mathrm{rank}\,B-\mathrm{rank}(BC) \tag{2.13.6c}$$

という“劣モジュラー不等式”と呼ばれる不等式に似た形をしたものがある．簡潔な証明をするにはやや準備が要るので，証明は 6.4 節でおこなう．(2.13.6)の $B=I_m$ の特別な場合として，

(vii)　A が $l\times m$ 型，B が $m\times n$ 型のとき，

$$\mathrm{rank}\,A+\mathrm{rank}\,B\leqq \mathrm{rank}(AB)+m. \tag{2.13.7}$$

2.13.2　項別階数

行列 A の階数は，0 にならないような小“行列式”の最大次数として定義されるが，この定義における“行列式”を“行列式の表式(2.6.1)における項”で置き換えて得られる概念が A の**項別階数**(term rank)で，t-rank A と記される．すなわち，A の行 $i_1,\cdots,i_r$，列 $j_1,\cdots,j_r$ (順序も意味がある)を $a_{i_1j_1}\neq 0,\cdots,a_{i_rj_r}\neq 0$ となるように選ぶことができるような最大の r を，A の項別階数と呼ぶ．定義から直ちに明らかなように，

$$\text{t-rank}\,A\geqq \mathrm{rank}\,A \tag{2.13.8}$$

である．A の“非零”要素の間に代数的関係が存在しないときには，t-rank A $=$rank A である．

項別階数は行列の要素が零であるか否かということのみに関係した概念である．上の定義をさらに言い換えれば，

　A の各行・各列からたかだか 1 個の非零要素を選ぶという選び方をして，
　選ぶことができる非零要素の最大数

ということにもなる．

注意 2.13.2　グラフ理論の用語(→補章 A.2 節)を用いれば，さらに，次のようにも述べられる．A の行に対応した m 個の頂点と列に対応した n 個の頂点とをもち，前者の第 i 頂点と後者の第 j 頂点とが，$a_{ij}\neq 0$ であるとき，そしてそのときに限り，辺で結ばれているような“2 部グラフ(bipartite graph)”における“最大マッチング(maximum matching)”の大きさ(そのマッチングを構成する辺の数)が A の項別階数である．

2.14　直交行列，ユニタリ行列，正規行列

本節で扱う行列は $K=\mathbb{C}$，すなわち複素数体の上の行列であるとする．もちろん，特別な場合として $K=\mathbb{R}$，すなわち実数体の上の行列も対象となる．

2.14.1 直 交 行 列

実数体 $\mathbb{R}$ の上の n 次正方行列 Q が

$$QQ^{\top}=I \tag{2.14.1}$$

を満たすとき，Q を**直交行列**(orthogonal matrix)と呼ぶ．逆行列の定義(→2.10.1 項)から，(2.14.1)は

$$Q^{\top}=Q^{-1} \quad \text{あるいは} \quad Q^{\top}Q=I \tag{2.14.2}$$

とも等価である．もちろん，直交行列は正則である．いまの場合，det Q が実数であるから，(2.8.16)により，

$$(\det Q)\cdot(\det Q^{\top})=(\det Q)^2=\det I=1$$

となる．したがって，直交行列 Q の行列式の値は

$$\det Q=1 \quad \text{あるいは} \quad -1 \tag{2.14.3}$$

である．直交行列 Q は det $Q=1$ のとき(広義の)**回転行列**とも呼ばれる．

(2.14.1)，(2.14.2)の関係式を要素ごとに書き下すと，

$$\begin{aligned}\boldsymbol{q}_{i\cdot}\cdot\boldsymbol{q}_{j\cdot}{}^{\top}=\boldsymbol{q}_{\cdot i}{}^{\top}\cdot\boldsymbol{q}_{\cdot j}&=\sum_{k=1}^{n}q_{ik}q_{jk}=\sum_{k=1}^{n}q_{ki}q_{kj}\\&=\begin{cases}1 & (i=j=1,\cdots,n\text{ のとき})\\0 & (i\neq j\text{ のとき})\end{cases}\end{aligned} \tag{2.14.4}$$

($\boldsymbol{q}_{\cdot i}$ は Q の第 i 列ベクトル(縦ベクトル)，$\boldsymbol{q}_{i\cdot}$ は第 i 行ベクトル(横ベクトル))となる．すなわち，直交行列の行ベクトルの集合および列ベクトルの集合は，それぞれ，“正規直交系”(→3.6 節)をなす．これが“直交”行列の名の由来である．

ベクトル $\boldsymbol{x}=[x_j]$の“長さ”を

$$\|\boldsymbol{x}\|=\sqrt{\boldsymbol{x}^{\top}\cdot\boldsymbol{x}}=\sqrt{\sum_{i=1}^{n}(x_i)^2} \tag{2.14.5}$$

で定義すると(より一般的な“ノルム”については第 3 章で扱う)，Q が直交行列のとき

$$\|Q\boldsymbol{x}\|=\|\boldsymbol{x}\| \tag{2.14.6}$$

である．実際，(2.14.2)，(2.14.5)より，

$$\|Q\boldsymbol{x}\|^2=(Q\boldsymbol{x})^{\top}\cdot Q\boldsymbol{x}=\boldsymbol{x}^{\top}Q^{\top}Q\boldsymbol{x}=\boldsymbol{x}^{\top}I\boldsymbol{x}=\boldsymbol{x}^{\top}\cdot\boldsymbol{x}=\|\boldsymbol{x}\|^2$$

である．

定義(2.14.1)と(2.14.2)とから明らかなように，Q が直交行列ならその転置

$Q^\top$ も直交行列である．

Q_1, Q_2 が共に直交行列なら Q_1Q_2 も直交行列である．実際，

$$Q_1{}^\top Q_1 = I, \quad Q_2{}^\top Q_2 = I$$

ならば

$$(Q_1Q_2)^\top(Q_1Q_2) = (Q_2{}^\top Q_1{}^\top)(Q_1Q_2) = Q_2{}^\top(Q_1{}^\top Q_1)Q_2 = Q_2{}^\top Q_2 = I$$

である．3 個以上の積についても同様．

変数 t の微分可能な関数を要素とする行列 $Q(t)$ が t のすべての値に対して直交行列であるとき，すなわち，

$$Q(t)^\top \cdot Q(t) = I \tag{2.14.7}$$

であるときには，(2.14.7) の両辺を t で微分すると (→(2.12.9))，

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} Q(t)\right)^\top \cdot Q(t) + Q(t)^\top \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t} Q(t)\right) = O. \tag{2.14.8}$$

ゆえに

$$\left(\frac{\mathrm{d}}{\mathrm{d}t} Q(t)\right)^\top = -Q(t)^\top \cdot \left(\frac{\mathrm{d}}{\mathrm{d}t} Q(t)\right) \cdot Q(t)^\top \tag{2.14.9}$$

という関係式が得られる．さらに，(2.14.8) を

$$\left(Q(t)^\top \cdot \frac{\mathrm{d}}{\mathrm{d}t} Q(t)\right)^\top = -Q(t)^\top \cdot \frac{\mathrm{d}}{\mathrm{d}t} Q(t) \tag{2.14.10}$$

と書き換えれば，$Q(t)^\top \cdot \dfrac{\mathrm{d}}{\mathrm{d}t} Q(t)$ が反対称行列であるということもわかる．

A が反対称行列ならば，$\exp(A)$ (→例 2.12.2) は直交行列である．実際，

$$\begin{aligned}(\exp(A))^\top \cdot \exp(A) &= \exp(A^\top) \cdot \exp(A) = \exp(-A) \cdot \exp(A) \\ &= \exp(O) = I.\end{aligned} \tag{2.14.11}$$

応用上有用な特殊な形の直交行列の例を二三挙げよう．

例 2.14.1 ほとんど自明であるが，置換行列 (→2.3.2 項) は直交行列である．互換行列の行列式は -1 に等しいから，偶(奇)置換を表す置換行列の行列式は $1(-1)$ に等しい．また，

$$Q = \begin{bmatrix} \pm 1 & & & O \\ & \pm 1 & & \\ & & \ddots & \\ O & & & \pm 1 \end{bmatrix} \tag{2.14.12}$$

も自明な直交行列である．直交三角行列がこの形のものに限られることも(2.14.4)から明らかであろう．　□

例 2.14.2　次の形の n 次正方行列(上 Hessenberg 行列の形をしている)は直交行列である．

$$Q=\begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \cdots & \frac{1}{\sqrt{n}} \\ -\sqrt{\frac{n-1}{n}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \cdots & \frac{1}{\sqrt{n(n-1)}} \\ 0 & -\sqrt{\frac{n-2}{n-1}} & \frac{1}{\sqrt{(n-1)(n-2)}} & \cdots & \cdots & \frac{1}{\sqrt{(n-1)(n-2)}} \\ & & \vdots & \vdots & \vdots & \vdots \\ & & 0 & -\sqrt{\frac{2}{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ & O & & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \tag{2.14.13}$$

$QQ^{\top}=I$ であることが直接計算により容易に確かめられる．　□

例 2.14.3　2 次元および 3 次元の回転行列．θ を任意の実数として，

$$Q=\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{2.14.14}$$

が直交行列で $\det Q=1$ であること，また，2 次の直交行列はこの形のものと，この第 1 行の符号を変えたもののどちらかに限られることは直ちに確かめられる．θ,φ,ψ を任意の実数として，

$$Q=\begin{bmatrix} \cos\varphi\cdot\cos\theta\cdot\cos\psi-\sin\varphi\cdot\sin\psi & \sin\varphi\cdot\cos\theta\cdot\cos\psi+\cos\varphi\cdot\sin\psi & -\sin\theta\cdot\cos\psi \\ -\cos\varphi\cdot\cos\theta\cdot\sin\psi-\sin\varphi\cdot\cos\psi & -\sin\varphi\cdot\cos\theta\cdot\sin\psi+\cos\varphi\cdot\cos\psi & \sin\theta\cdot\sin\psi \\ \cos\varphi\cdot\sin\theta & \sin\varphi\cdot\sin\theta & \cos\theta \end{bmatrix} \tag{2.14.15}$$

が直交行列で $\det Q=1$ であることも直接計算により確かめられる．これらの行列は，2 次元あるいは 3 次元の直交直線座標系の回転を表す行列であり(→3.3 節)，(2.14.14)の θ は平面の中での回転角を，(2.14.15)の θ, φ, ψ はいわゆる "Euler(オイラー)角" を表している． □

例 2.14.4 Householder 行列．長さが 1 に等しい任意の縦ベクトル($n\times 1$ 型の行列) $\boldsymbol{u}$ ($\boldsymbol{u}^{\top}\cdot\boldsymbol{u}=1$) に対して，

$$H(\boldsymbol{u})=I_n-2\boldsymbol{u}\boldsymbol{u}^{\top} \tag{2.14.16}$$

で定義される n 次正方行列 $H(\boldsymbol{u})$ を，ベクトル $\boldsymbol{u}$ により定められる **Householder(ハウスホルダー)行列**という．定義から直ちにわかるように，Householder 行列は対称行列である：

$$H(\boldsymbol{u})^{\top}=H(\boldsymbol{u}). \tag{2.14.17}$$

$H(\boldsymbol{u})\cdot H(\boldsymbol{u})$ を計算すると，$\boldsymbol{u}^{\top}\boldsymbol{u}=1$ であるから，

$$\begin{aligned} H(\boldsymbol{u})\cdot H(\boldsymbol{u}) &= (I-2\boldsymbol{u}\boldsymbol{u}^{\top})(I-2\boldsymbol{u}\boldsymbol{u}^{\top})=I-4\boldsymbol{u}\boldsymbol{u}^{\top}+4\boldsymbol{u}(\boldsymbol{u}^{\top}\boldsymbol{u})\boldsymbol{u}^{\top} \\ &= I-4\boldsymbol{u}\boldsymbol{u}^{\top}+4\boldsymbol{u}\boldsymbol{u}^{\top}=I \end{aligned} \tag{2.14.18}$$

となる．(2.14.17)と(2.14.18)とから，Householder 行列が直交行列であることがわかる．また，

$$\det H(\boldsymbol{u})=-1 \tag{2.14.19}$$

である．(注意 2.10.4 の(2.10.9′)によって，$\det H(\boldsymbol{u})=(\det I)\cdot(1-2\boldsymbol{u}^{\top}I\boldsymbol{u})=1-2=-1$.)

ある縦ベクトル $\boldsymbol{x}$ に左から $H(\boldsymbol{u})$ を掛けると，

$$H(\boldsymbol{u})\cdot\boldsymbol{x}=\boldsymbol{x}-2(\boldsymbol{u}^{\top}\cdot\boldsymbol{x})\boldsymbol{u} \tag{2.14.20}$$

となるが，この式の右辺第 2 項は「$\boldsymbol{x}$ を $\boldsymbol{u}$ 方向に正射影したものの 2 倍を $\boldsymbol{x}$ から引いたもの」すなわち「$\boldsymbol{u}$ を法線とする(超)平面を鏡としたときの $\boldsymbol{x}$ の鏡像」という解釈ができる(図 2.14.1)(ベクトル解析に関する書物も参照されたい)．

そこで，$\|\boldsymbol{x}\|=\|\boldsymbol{y}\|$，$\boldsymbol{x}\neq\boldsymbol{y}$ であるような 2 本のベクトル $\boldsymbol{x}, \boldsymbol{y}$ に対して，$\boldsymbol{u}=(\boldsymbol{x}-\boldsymbol{y})/\|\boldsymbol{x}-\boldsymbol{y}\|$ とおけば，

$$\boldsymbol{y}=H(\boldsymbol{u})\boldsymbol{x}$$

となる．(これは(2.14.20)に代入して直接計算すれば確かめられる．(2.14.33)～(2.14.35)も見よ．)

なお，実際に数値計算などをするときには，Householder 行列を "行列" と

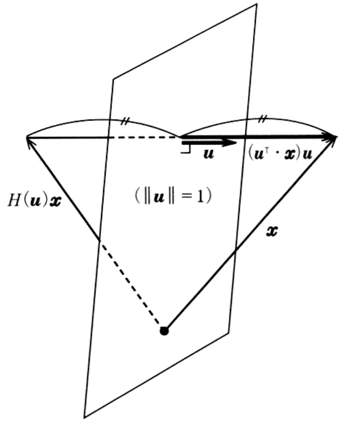


図 2.14.1　$H(\boldsymbol{u})\cdot\boldsymbol{x}$

して表現して，$H(\boldsymbol{u})\boldsymbol{z}$ を行列 $H(\boldsymbol{u})$ とベクトル $\boldsymbol{z}$ の積として計算するようなことは，決しておこなってはならない．$H(\boldsymbol{u})$ の情報は $\boldsymbol{u}$ にすべて収められているので，$H(\boldsymbol{u})$ は $\boldsymbol{u}$ で表し，$H(\boldsymbol{u})\boldsymbol{z}$ は $\boldsymbol{z}-2(\boldsymbol{u}^{\mathsf{T}}\cdot\boldsymbol{z})\boldsymbol{u}$ と計算しなければならない．(より詳しくは数値計算法の書物を参照のこと．)　□

2.14.2　ユニタリ行列

複素数体 $\mathbb{C}$ の上の n 次正方行列 U が

$$U^{\mathsf{H}}U=I \tag{2.14.21}$$

を満たすとき，U を**ユニタリ行列**(unitary matrix)と呼ぶ．実数を要素とする直交行列を複素数体の上の行列とみなせば，それはユニタリ行列である．ユニタリ行列は，直交行列の複素版であって，直交行列の有していた諸性質はユニタリ行列に拡張される．

まず，U がユニタリ行列ならば

$$U^{\mathsf{H}}=U^{-1},\quad UU^{\mathsf{H}}=I \tag{2.14.22}$$

である．また，$(\det U)\cdot(\det U^{\mathsf{H}})=|\det U|^2=1$ より，

$$|\det U|=1,\quad \text{すなわち}\quad \det U=\mathrm{e}^{\mathrm{i}\theta}\quad (\theta\ \text{は実数}) \tag{2.14.23}$$

である．U の第 i 列(縦)ベクトルを $\boldsymbol{u}_{\cdot i}$，第 i 行(横)ベクトルを $\boldsymbol{u}_{i\cdot}$ と記すと，

$$\boldsymbol{u}_{i\cdot}\cdot\boldsymbol{u}_{j\cdot}{}^{\mathsf{H}}=\boldsymbol{u}_{\cdot i}{}^{\mathsf{H}}\cdot\boldsymbol{u}_{\cdot j}=\sum_{k=1}^{n}u_{ik}\overline{u_{jk}}=\sum_{k=1}^{n}\overline{u_{ki}}u_{kj}$$
$$=\begin{cases}1 & (i=j\text{ のとき})\\ 0 & (i\neq j\text{ のとき})\end{cases} \tag{2.14.24}$$

となる．

成分が複素数のベクトル $\boldsymbol{x}$ の“長さ”を

$$\|\boldsymbol{x}\|=\sqrt{\boldsymbol{x}^{\mathsf{H}}\cdot\boldsymbol{x}}=\sqrt{\sum_{i=1}^{n}|x_i|^2} \tag{2.14.25}$$

で定義すると，U がユニタリ行列のとき，

$$\|U\boldsymbol{x}\|=\|\boldsymbol{x}\| \tag{2.14.26}$$

となる．

U_1, U_2 が共にユニタリ行列なら U_1U_2 もユニタリ行列である．U がユニタリ行列なら $U^{\mathsf{H}}, U^{\mathsf{T}}, \overline{U}$ もユニタリ行列である．

t をパラメタにもつユニタリ行列 $U(t)$ に対して，

$$\left(U(t)^{\mathsf{H}}\cdot\frac{\mathrm{d}}{\mathrm{d}t}U(t)\right)^{\mathsf{H}}=-U(t)^{\mathsf{H}}\cdot\frac{\mathrm{d}}{\mathrm{d}t}U(t) \tag{2.14.27}$$

となる．すなわち，$U(t)^{\mathsf{H}}\cdot\dfrac{\mathrm{d}}{\mathrm{d}t}U(t)$ は歪 Hermite 行列である．(証明は(2.14.10)のときと同様．)

A が歪 Hermite 行列ならば，$\exp(A)$ (→例 2.12.2) はユニタリ行列である．(証明は(2.14.11)にならえば容易．)

例 2.14.5 ユニタリな対角行列は $\mathrm{diag}[\mathrm{e}^{\mathrm{i}\sigma_1},\mathrm{e}^{\mathrm{i}\sigma_2},\cdots,\mathrm{e}^{\mathrm{i}\sigma_n}]$ (σ_i は実数) という形をしている．(2.14.24)によれば，ユニタリ三角行列はこの形のものに限られる．

2 次のユニタリ行列は，実数 ρ,σ,τ,θ を用いて

$$U=\begin{bmatrix}\mathrm{e}^{\mathrm{i}(\rho+\sigma)}\cos\theta & \mathrm{e}^{\mathrm{i}(\rho+\tau)}\sin\theta\\ -\mathrm{e}^{\mathrm{i}(\rho-\tau)}\sin\theta & \mathrm{e}^{\mathrm{i}(\rho-\sigma)}\cos\theta\end{bmatrix} \tag{2.14.28}$$

という形(あるいは第 1 行の符号を変えたもの)に表される． □

例 2.14.6 Householder 行列の複素版は，$\boldsymbol{u}^{\mathsf{H}}\cdot\boldsymbol{u}=1$ であるような(すなわち，“長さ”が 1 の)任意の $n\times 1$ 型行列(縦ベクトル) $\boldsymbol{u}$ に対して，

$$H(\boldsymbol{u})=I_n-2\boldsymbol{u}\boldsymbol{u}^{\mathsf{H}} \tag{2.14.29}$$

で定義される．$H(\boldsymbol{u})$ が Hermite 行列でユニタリ行列であることも明らかであ

る．

$$H(\boldsymbol{u})^{\mathrm{H}}=H(\boldsymbol{u}),\quad H(\boldsymbol{u})H(\boldsymbol{u})=I. \tag{2.14.30}$$

また，

$$\det H(\boldsymbol{u})=-1 \tag{2.14.31}$$

であることも実数のときと同様．さらに，複素数ベクトル $\boldsymbol{x}$ に対しては，

$$H(\boldsymbol{u})\boldsymbol{x}=\boldsymbol{x}-2(\boldsymbol{u}^{\mathrm{H}}\cdot\boldsymbol{x})\boldsymbol{u} \tag{2.14.32}$$

と表される．いま，

$$\|\boldsymbol{x}\|=\|\boldsymbol{y}\|,\quad \boldsymbol{x}^{\mathrm{H}}\cdot\boldsymbol{y}=\boldsymbol{y}^{\mathrm{H}}\cdot\boldsymbol{x}\quad (\text{すなわち } \boldsymbol{x}^{\mathrm{H}}\cdot\boldsymbol{y} \text{ が実数}),\ \boldsymbol{x}\neq\boldsymbol{y} \tag{2.14.33}$$

を満たす2本のベクトル $\boldsymbol{x},\boldsymbol{y}$ に対して，

$$\boldsymbol{u}=\frac{\boldsymbol{x}-\boldsymbol{y}}{\|\boldsymbol{x}-\boldsymbol{y}\|} \tag{2.14.34}$$

とおけば，

$$\begin{aligned} H(\boldsymbol{u})\boldsymbol{x}&=\boldsymbol{x}-2\frac{(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{H}}\cdot\boldsymbol{x}}{(\boldsymbol{x}-\boldsymbol{y})^{\mathrm{H}}\cdot(\boldsymbol{x}-\boldsymbol{y})}(\boldsymbol{x}-\boldsymbol{y})\\ &=\boldsymbol{x}-\frac{2(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}-\boldsymbol{y}^{\mathrm{H}}\boldsymbol{x})}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}+\boldsymbol{y}^{\mathrm{H}}\boldsymbol{y}-\boldsymbol{y}^{\mathrm{H}}\boldsymbol{x}-\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}}(\boldsymbol{x}-\boldsymbol{y})\\ &=\boldsymbol{x}-\frac{2(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}-\boldsymbol{y}^{\mathrm{H}}\boldsymbol{x})}{2(\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}-\boldsymbol{y}^{\mathrm{H}}\boldsymbol{x})}(\boldsymbol{x}-\boldsymbol{y})=\boldsymbol{y} \end{aligned} \tag{2.14.35}$$

となる．□

例 2.14.7　複素数体の上で $\zeta=\exp(2\pi\mathrm{i}/n)$ とする．（一般の体の上でも，1の“原始 n 乗根”（の一つ）を ζ とすれば以下の話は成り立つ．1の原始 n 乗根とは，$\zeta^n=1, \zeta^m\neq 1\,(m=1,2,\cdots,n-1)$ を満たす ζ のことである．）

n 次正方行列 $Z=[z_{ij}]$ を

$$z_{ij}=\frac{1}{\sqrt{n}}\zeta^{(i-1)\cdot(j-1)}\qquad (i,j=1,\cdots,n) \tag{2.14.36}$$

で定義する．すると

$$\begin{aligned} \sum_{k=1}^{n}\overline{z_{ki}}z_{kj}&=\frac{1}{n}\sum_{k=1}^{n}\zeta^{(k-1)\cdot(j-i)}=\frac{1}{n}\sum_{k=0}^{n-1}\zeta^{k\cdot(j-i)}=\frac{1}{n}\sum_{k=0}^{n-1}(\zeta^{j-i})^k\\ &=\begin{cases} 1 & (i=j \text{ のとき})\\ \dfrac{1}{n}\dfrac{1-\zeta^{n\cdot(j-i)}}{1-\zeta^{j-i}}=0 & (i\neq j \text{ のとき}) \end{cases} \end{aligned} \tag{2.14.37}$$

である(一般の体の上では $\overline{\zeta}=1/\zeta$ とする)．したがって，

$$Z^{\mathrm{H}}Z=I_n \tag{2.14.38}$$

となる．すなわち，Zは(対称な)ユニタリ行列である．

n 次元縦ベクトル $\boldsymbol{x}$ に対して，$\boldsymbol{y}=Z\boldsymbol{x}$ は $\boldsymbol{x}$ の**離散 Fourier**(フーリエ)**変換**(discrete Fourier transform)と呼ばれる．$\boldsymbol{x}$ から $\boldsymbol{y}$ への写像 Z のことを**離散 Fourier 変換**(discrete Fourier transformation)，$Z^{-1}=Z^{\mathrm{H}}$ の表す変換を**離散 Fourier 逆変換**(inverse discrete Fourier transformation)という．これらは信号処理技術などにおいて重要な役割を果たしている．特に，$\boldsymbol{y}=Z\boldsymbol{x}$ の計算を n^2 回の乗算をせずに $n\log_2 n$ に比例する程度の回数の乗算で済ませてしまう計算法(“高速 Fourier 変換”(Fast Fourier Transformation, FFT)と呼ばれる)が知られるようになってから，離散 Fourier 変換は実用的な数値計算においても理論的な計算複雑度の研究においても活発に利用されるようになっている．　□

2.14.3 正規行列

複素数体上での正方行列 A は

$$A^{\mathrm{H}}A=AA^{\mathrm{H}} \tag{2.14.39}$$

を満たすとき**正規行列**(normal matrix)と呼ばれる．A^{H} と A が可換であることは，A の固有値問題，Jordan 標準形に関連して重要な性質である(→5.2 節，6.8 節)．

Hermite 行列(特殊な場合として，実対称行列)，歪 Hermite 行列(特殊な場合として，実反対称行列)，ユニタリ行列(特殊な場合として，(実)直交行列)などが正規行列であることは明らかであろう．

2.15 基本的な変形

行列あるいはそれにより表される関係式の形を見やすくしたり扱いやすくしたりするために役に立ついくつかの基本的な方法・手続きを以下に示す．

2.15.1 基本変形

2.3.2 項で定義した基本行列 $P(i_0,j_0)$，$T(i_0\,;\,t)\,(t\neq 0)$，$W(i_0,j_0\,;\,t)$を，行列 A の左から掛けることを A に対する**基本行演算**(elementary row operation)，

右から掛けることを**基本列演算**(elementary column operation)という．すなわち，基本行(列)演算とは以下の演算である：

(i) 行(列) i_0 と j_0 とを交換する，

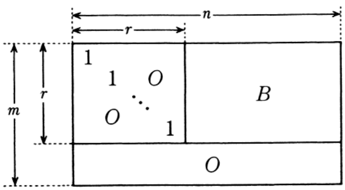
(ii) 行(列) i_0 を t $(\neq 0)$ 倍する，

(iii) 行 j_0 (列 i_0) を t 倍して行 i_0 (列 j_0) に加える．

基本行演算，基本列演算により行列を変形することを，行列の**基本変形**(elementary transformation)という．

基本行列は正則であるから(→(2.6.14)，(2.6.15)，(2.6.16))，基本変形により行列の階数は変化しない(→(2.13.3))．

$m\times n$ 型の行列 A は，列の置換と行に関する基本変形によって，

$\tilde{A}=$  (2.15.1)

という形にすることができる．ここで，$r=\mathrm{rank}\,A$ である．($\mathrm{rank}\,\tilde{A}=r$ であることは明らかであり，$\mathrm{rank}\,A=\mathrm{rank}\,\tilde{A}$ であるから．)それには次の手続きによればよい．この手続きは，与えられた行列 A の階数を定めるための具体的な手順を述べていることにもなる．(しかし，数値計算に伴う丸め誤差があるような場合には，階数を定めることには原理的な困難がある．行列のある要素が数値的に0に等しいかどうかを判別することが，丸め誤差が存在する環境のもとでは，原理的に困難であるからである．)

[1] $A^{(1)}:=A$, $k:=1$ とおき，[2]へ進む．

[2] $A^{(k)}$ の列 $k,k+1,\cdots,n$ の中で，行 $k,k+1,\cdots,m$ のいずれかに非零要素をもつものを探す．

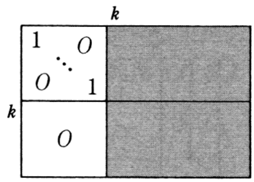
[2.1] 非零要素 $a_{i_0j_0}^{(k)}\neq 0$ が見出だされたら，$A^{(k)\prime}:=P(i_0,k)A^{(k)}P(k,j_0)$ を作り，[3]へ進む．

[2.2] そこに非零要素が見出だされなかったら($k=n+1$ あるいは $k=m+1$ となったときも，形式的にはこの場合であるとする)，$r:=k-1$,

$\tilde{A} := A^{(k)}$ として，手続きを終了する．

[3]　$A^{(k)\prime}$ の左から $T(k\,;1/a_{kk}^{(k)\prime})$ を掛け（すなわち，$A^{(k)\prime}$ の行 k を $a_{kk}^{(k)\prime} = a_{i_0 j_0}^{(k)}$ で割って (k,k) 要素を1に等しくし），さらに左から $W(i,k\,; -a_{ik}^{(k)\prime})$ を $i:=1,2,\cdots,k-1,k+1,\cdots,m$ に対して次々と掛け（すなわち，行 k 以外のすべての行から行 k を $a_{ik}^{(k)\prime}$ 倍したものを引き），このようにして得られる行列を $A^{(k+1)}$ とし，$k+1$ をあらためて k とおいて，[2]へ戻る．

上の手続きに従えば，$A^{(k)}$ は

$A^{(k)} =$ 　(2.15.2)

という形をしており，[3]において，列 k は，行 k の要素が1でそれ以外の行の要素が0であるような列に変形される（列 $1,\cdots,k-1$ は[3]で変化しない）．

この基本変形をまとめて行列の間の関係式として書き表せば，S をある m 次正則行列，P を n 次置換行列として，

$$SAP = \tilde{A} \tag{2.15.3}$$

となる．（実際には，$S^{(1)} = I_m$，$P^{(1)} = I_n$ から出発して $A^{(k)}$ に施す基本行演算と同じものを $S^{(k)}$ に，$A^{(k)}$ の列の互換と同じ互換を $P^{(k)}$ の列に，それぞれ施して $S^{(k+1)}, P^{(k+1)}$ を作っていけば，$S = S^{(r+1)}$，$P = P^{(r+1)}$ が得られる．）

注意 2.15.1　以上の話は行列の行と列の役割を入れ換えても成立することは明らかであろう．

注意 2.15.2　A が正則な正方行列のときに，$\tilde{A} = I_n$ であるから，(2.15.3)より，

$$A = S^{-1}P^{-1} \quad \text{あるいは} \quad A^{-1} = PS \tag{2.15.4}$$

である．そこで，A に対して基本変形を繰り返し施すとき，それと同じ行演算を I_n に対して施していけば，A から $\tilde{A} = I_n$ が得られたとき，I_n からは S が得られることになる．A の列に対して適用したのに対応する置換を S の行に対して適用すれば A^{-1} が得られる．（A の列に対して適用した互換を，逆順に S の行に対して適用すればよい．）

注意 2.15.3　基本行列，基本行(列)演算，基本変形などは，ブロック行列(→2.4.1 項)の場合にも，各ブロックの型に注意しさえすれば，自然に拡張することができる．このとき，対角要素に対しては正方ブロックが，非零要素に対しては正則ブロックが，それぞれ対応する．

注意 2.15.4　(2.15.1)の $\tilde{A}$ に対して，さらに基本**列**演算を施して，

$$SAT = \tilde{\tilde{A}} = \overset{\quad n\quad}{\underset{m}{\left[\begin{array}{c|c} \begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix} \ (r \times r) & O \\ \hline O & O \end{array}\right]}} \tag{2.15.5}$$

という形にすることも可能である(→6.4 節)．ここで，T は次の n 次正則行列である．

$$T = P\cdot\left[\begin{array}{c|c} I_r & -B \\ \hline O & I_{n-r} \end{array}\right]. \tag{2.15.6}$$

注意 2.15.5　詳細は数値計算の専門書で「線形計算」を扱っているものに譲るべきことであるが，上記の手続き[1]〜[3]を実行するのに要する演算(加減算，乗除算，大小比較・符号判定，入れ換え)の回数に，ちょっと注目してみよう．$A^{(k)}$ から $A^{(k+1)}$ を作るには，[2.1]においてたかだか $(n-k+1)\cdot(m-k+1)$ 回の符号判定と $m+(n-k+1)$ 回の入れ換えを，[3]においてたかだか 1 回の除算と $m(n-k)$ 回の乗算と $(m-1)\cdot(n-k)$ 回の減算とを，それぞれおこなえばよい．そこで，A から $\tilde{A}$ を得るまでには，[2.1]が $k=1,\cdots,r+1$ に対して，[3]が $k=1,\cdots,r$ に対して，それぞれ，実行されるから，必要とされる演算の総数は，たかだか

$$\sum_{k=1}^{r+1}(n-k+1)(m-k+1)=(r+1)\left(nm-\frac{n+m}{2}r+\frac{1}{3}r\left(r+\frac{1}{2}\right)\right) \tag{2.15.7}$$

回の符号判定と

$$\sum_{k=1}^{r+1}(m+(n-k+1))=(r+1)\left(n+m-\frac{r}{2}\right) \tag{2.15.8}$$

回の入れ換えと

$$\sum_{k=1}^{r}1=r \tag{2.15.9}$$

回の除算と

$$\sum_{k=1}^{r}m(n-k)=rm\left(n-\frac{r+1}{2}\right) \tag{2.15.10}$$

回の乗算と，それに，

$$\sum_{k=1}^{r}(m-1)(n-k)=r(m-1)\left(n-\frac{r+1}{2}\right) \tag{2.15.11}$$

回の減算とである．n, m, r が同程度の大きさならば，これらの演算の中で大勢を占めるのは，ほぼ $n^3/3$ 回程度の符号判定と $n^3/2$ 回程度の乗算および(加)減算とである．S も計算しようとすると，さらに，ほぼ $n^3/2$ 回程度の乗算および加減算が必要となる(詳細は省く)．

注意 2.15.6　上記の手続きの中で，[3]に登場する $a_{kk}^{(k)\prime}=a_{i_0j_0}^{(k)}$ は，通常，**軸要素** (pivot) と呼ばれる．

(2.15.1)，(2.15.3)をさらに詳しく眺めてみよう．(2.15.3)の S は基本行列の積であるが，(2.3.11)，(2.3.12)からもわかるように，置換行列の部分を右に寄せてしまうことができる．換言すれば，A から $\tilde{A}$ を作る手続きの中でおこなった"行の入れ換え"をあらかじめ A に施しておくことができる．

ここでは，A は既にそのような行の入れ換えをおこなってあるものとしよう．列についても同様とする．すると，上記の手続きの中で実質的に残る計算は，[3]だけになる．[3]において，軸要素は最初の r 行から選ばれているので，行 $r+1, \cdots, m$ は，他の行の何倍かを加減されるだけで，それらの行を何倍かして他の行に加減するということはなされない．そこで，(2.15.3)の形をより具体的に書き表すと，

$$\begin{array}{c|c|c|}
 & 1\cdots r & r+1\cdots m \\ \hline
1\cdots r & S_{11} & O \\ \hline
r+1\cdots m & S_{21} & \begin{smallmatrix}1 & & O\\ & \ddots & \\ O & & 1\end{smallmatrix} \\ \hline
\end{array}
\cdot
\begin{array}{|c|c|}
1\cdots r & r+1\cdots n \\ \hline
A_{11} & A_{12} \\ \hline
A_{21} & A_{22} \\ \hline
\end{array}
=
\begin{array}{c|c|c|}
 & 1\cdots r & r+1\cdots n \\ \hline
1\cdots r & \begin{smallmatrix}1 & & O\\ & \ddots & \\ O & & 1\end{smallmatrix} & B \\ \hline
r+1\cdots m & \multicolumn{2}{c|}{O} \\ \hline
\end{array} \tag{2.15.12}$$

という形になる．ブロックごとの関係を書くと，

$$\begin{aligned} &S_{11}A_{11}=I_r, \quad S_{11}A_{12}=B, \\ &S_{21}\cdot[A_{11}, A_{12}]+[A_{21}, A_{22}]=O_{m-r,n} \end{aligned} \tag{2.15.13}$$

となる．

これから次のことがわかる．(なお，(2.10.9)も参照のこと．)

(i)　A_{11} は A の階数 r に等しい次数の(すなわち最大次数の)正則部分行列であり，$S_{11}=A_{11}^{-1}$ である．

(ii)　$B=A_{11}^{-1}A_{12}$．

(iii)　$S_{21}=-A_{21}A_{11}^{-1}$．

(iv) $A_{22}=A_{21}A_{11}{}^{-1}A_{12}$；すなわち，$A$ の階数が r であると，A_{11}, A_{12}, A_{21} から A_{22} の部分が定まってしまう．

注意 2.15.7 2.15.1 項の行に関する基本変形(i)，(ii)，(iii)を $m\times n$ 型の行列 A に施すとき，A の r 次の小行列式から作られる $\binom{m}{r}\times\binom{n}{r}$ 型の行列 $A^{<r>}$ (→ 2.7.1 項)も，行に関する基本変形をいくつか施されたものに変形される．実際，A に(i)の形の基本変形を施す，すなわち A の行 i_0 と j_0 を交換すると，$A^{<r>}$ のいくつかの行の符号が変わったり順序が入れ換わったりする．($|A|^{i_1\cdots i_r}_{j_1\cdots j_r}$ の $\{i_1,\cdots,i_r\}$ の中に i_0 と j_0 が共に含まれると，その符号が変わる．i_0 と j_0 の一方，たとえば i_0，だけが含まれるようなものは，i_0 のところだけを j_0 におきかえ $i_1<i_2<\cdots<i_r$ という約束に従って順序を入れ換えたものと入れ換わる(符号が変わることもある)．i_0, j_0 のどちらも含まれないときは変化しない．) A に(ii)の型の変形を施すと，すなわち A の行 i_0 を t 倍すると，$A^{<r>}$ の要素 $|A|^{i_1\cdots i_r}_{j_1\cdots j_r}$ で $\{i_1,\cdots,i_r\}$ の中に i_0 が含まれるようなものが一斉に t 倍される，すなわち $A^{<r>}$ のいくつかの行が t 倍される．A に(iii)の形の変形を施すと，すなわち行 j_0 を t 倍して行 i_0 に加えると，$A^{<r>}$ のいくつかの行が t 倍されて他のいくつかの行に加えられる．($\{i_1,\cdots,i_r\}$ の中に i_0 と j_0 が共に含まれるような $A^{<r>}$ の行は変化しない(行列式の基本的性質：注意 2.6.2(ハ).) j_0 のみが含まれているような $A^{<r>}$ の行は変化しない．i_0 のみが含まれているような $A^{<r>}$ の行には，その $\{i_1,\cdots,i_r\}$ の中の i_0 のところを j_0 でおきかえ(て必要なら順序を入れ換え符号を変え)たものを t 倍したものが加えられる．

基本列変形についても同様．

基本変形で行列の階数は変化しない．rank $A=s$ であるとき(2.15.5)の形の $\tilde{\tilde{A}}$ を作ると，rank $\tilde{\tilde{A}}=s$ である．このとき rank $\tilde{\tilde{A}}^{<r>}=$rank $A^{<r>}$ である．ところで，$r>s$ の $A^{<r>}$ の(そして $\tilde{\tilde{A}}^{<r>}$ の)すべての要素は 0 である．$r=s$ なら $|\tilde{\tilde{A}}^{<r>}|^{1\cdots r}_{1\cdots r}=1$ で他の要素はみな 0 であるから，rank $\tilde{\tilde{A}}^{<r>}=$rank $A^{<r>}=1$ である．$r<s$ ならば，$|\tilde{\tilde{A}}|^{i_1\cdots i_r}_{i_1\cdots i_r}=1$ ($i_1<\cdots<i_r$, $i_k\leqq s$)で，他の要素は 0 であるから，rank $\tilde{\tilde{A}}^{<r>}=$rank $A^{<r>}=\binom{s}{r}$ である．

このようにして，A の小行列式からなる $A^{<r>}$ の階数が具体的に定められる．

2.15.2 枢軸変換

n 個の変数 $x_1,\cdots,x_n$ がまったく独立に自由にその値を取るのではなくて，ある与えられた $m\times n$ 型の行列 $A=[a_{ij}]$ に対して，

$$\sum_{j=1}^{n} a_{ij}x_j=0 \qquad (i=1,\cdots,m) \tag{2.15.14}$$

という関係を満たす範囲で自由に値を取ることができるというような場合のこと

を考える．

変数を $\boldsymbol{x}=[x_1,\cdots,x_n]^{\top}$ という n 次の縦ベクトル($n\times1$ 型行列)にまとめて表せば，(2.15.14)は，

$$A\boldsymbol{x}=\mathbf{0} \tag{2.15.14$'$}$$

($\mathbf{0}$ は m 次零縦ベクトル($=O_{m,1}$))となる．任意の m 次正則行列 S に対して，(2.15.14′)より，

$$(SA)\boldsymbol{x}=\mathbf{0} \tag{2.15.15}$$

が得られるし，また(2.15.15)から(2.15.14′)が得られるから，(2.15.14)は適当に S を選んで“見やすい形”にして扱うのがよい．

本節前項(2.15.1 基本変形)において調べたように，rank $A=r$ のとき，S を適当に選べば，(2.15.15)は，

$$x_i+\sum_{j\in J}b_{ij}x_j=0 \qquad (i\in I) \tag{2.15.16}$$

のように表せる．ここで，I は(2.15.1)において $\tilde{A}$ の列の左側の r 本に対応する変数(列)番号の集合($\tilde{A}$ の列は A の列を置換したものになっていることに注意！)，J は右側の $n-r$ 本の列に対応する変数(列)番号の集合とする．

$$\begin{aligned}&|I|=r, \quad |J|=n-r,\\ &I\cap J=\varnothing, \quad I\cup J=N\equiv\{1,2,\cdots,n\}.\end{aligned} \tag{2.15.17}$$

(2.15.16)は「変数 $x_i\,(i\in I)$ を変数 $x_j\,(j\in J)$ で陽に表す式」とみなせる．このような表現形式を **基底形式** (basic form)，変数 $x_i\,(i\in I)$ のことを **基底変数** (basic variables)，$x_j\,(j\in J)$ のことを**非基底変数**(nonbasic variables)と呼ぶ．(I あるいは(I,J)のことを**基底**(basis)と呼ぶこともある．第3章ベクトル空間に現れる“基底”との間には「A の列ベクトルの張る空間において，I に対応する列ベクトルの集合が一つの基底をなす」という意味の関係がある．)これらの用語は主として数理計画法(特に，線形計画法)において用いられる用語である．

2.1節の注意2.1.1で述べたように，変数の番号づけには特に意味がないことが多いから，N,I,J は(特に順序はついていない)集合であるとみなして，以下の議論を進める．

(2.15.16)に現れる係数 b_{ij} を要素とする $r\times(n-r)$ 型行列を $B: I\times J\to K$ とする．$b_{i_0j_0}\neq0$ であるような $i_0\in I$，$j_0\in J$ を任意に選んで，i_0 に対応する式を $b_{i_0j_0}$ で割ると，

$$x_{j_0}+(1/b_{i_0j_0})x_{i_0}+\sum_{j\in J-\{j_0\}}(b_{i_0j}/b_{i_0j_0})x_j=0 \qquad (2.15.18)$$

となる．これを，x_{j_0} を添字が $\{i_0\}\cup(J-\{j_0\})$ に含まれる変数で表す式とみなして，(2.15.16)の $i=i_0$ 以外の式に代入すると，

$$x_i-(b_{ij_0}/b_{i_0j_0})x_{i_0}+\sum_{j\in J-\{j_0\}}(b_{ij}-b_{ij_0}b_{i_0j}/b_{i_0j_0})x_j=0$$
$$(i\in I-\{i_0\}) \qquad (2.15.19)$$

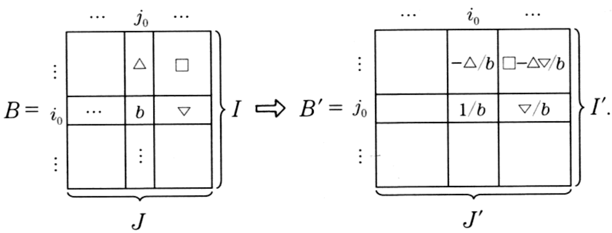
を得る．(2.15.18)と(2.15.19)をまとめると，

$$\begin{aligned}&x_i+\sum_{j\in J'}b'_{ij}x_j=0 \qquad (i\in I'),\\ &I'=(I-\{i_0\})\cup\{j_0\}, \qquad J'=(J-\{j_0\})\cup\{i_0\}\end{aligned} \qquad (2.15.20)$$

という形の別の基底形式が得られる(I と I'，J と J' では i_0 と j_0 が入れ換わっているだけ)．ここで，

$$\begin{aligned}&b'_{j_0i_0}=1/b_{i_0j_0},\\ &b'_{j_0j}=b_{i_0j}/b_{i_0j_0} \qquad (j\in J-\{j_0\}=J'-\{i_0\}),\\ &b'_{ii_0}=-b_{ij_0}/b_{i_0j_0} \qquad (i\in I-\{i_0\}=I'-\{j_0\}),\\ &b'_{ij}=b_{ij}-b_{ij_0}b_{i_0j}/b_{i_0j_0} \qquad \begin{pmatrix}i\in I-\{i_0\}=I'-\{j_0\}\\ j\in J-\{j_0\}=J'-\{i_0\}\end{pmatrix}\end{aligned} \qquad (2.15.21)$$

である．図的に覚えやすく表現すると



(2.15.22)

$B=[b_{ij}]$から $B'=[b'_{ij}]$へのこの変換を，(i_0,j_0)(または $b_{i_0j_0}$)を**軸**(あるいは**枢軸**；pivot)とする**枢軸変換**(pivotal transformation)と呼ぶ．

より一般に，二つの異なる基底(I,J)，(I',J')に対して

$$I_0=I-I'\overset{*}{=}J'-J, \quad J_0=J-J'\overset{*}{=}I'-I \qquad (2.15.23)$$

とおき((2.15.17)により，＊印のついた等式が成り立つこと，また基底が異なるから $|I_0|=|J_0|\neq 0$ であることに注意)，対応する行列 B,B' を次のようにブロックに分割して表す．

$$B = \begin{array}{c|c|c|} & J_0 & J-J_0(=J'-I_0) \\ \hline I_0 & B_{11} & B_{12} \\ \hline \begin{array}{c} I-I_0 \\ (=I'-J_0) \end{array} & B_{21} & B_{22} \\ \hline \end{array}, \qquad B' = \begin{array}{c|c|c|} & I_0 & J'-I_0(=J-J_0) \\ \hline J_0 & B'_{11} & B'_{12} \\ \hline \begin{array}{c} I'-J_0 \\ (=I-I_0) \end{array} & B'_{21} & B'_{22} \\ \hline \end{array}. \tag{2.15.24}$$

すると，B_{11} は正則で，

$$\begin{aligned} B'_{11} &= B_{11}{}^{-1}, \\ B'_{12} &= B_{11}{}^{-1}\cdot B_{12}, \\ B'_{21} &= -B_{21}\cdot B_{11}{}^{-1}, \\ B'_{22} &= B_{22}-B_{21}\cdot B_{11}{}^{-1}\cdot B_{12} \end{aligned} \tag{2.15.25}$$

であることがわかる．((2.15.18)～(2.15.21)の式変形を，要素 b_{ij} についてでなく，ブロック行列 $B_{11},\cdots$ について書き直せばよい．)

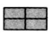
このBから B' への変換も (I_0, J_0) (または B_{11}) を枢軸とする**枢軸変換**と呼ばれる．

また，枢軸変換によって互いに結ばれる行列は，**組合せ論的に等価**(combinatorially equivalent)であると呼ばれることがある．

注意 2.15.8　(2.15.24)の B_{11} が正則でなければならないことは，基底の意味からも明らかであろうが，より形式的には，B に対応する $x_1,\cdots,x_n$ の間の関係式

$$\underbrace{\begin{array}{|c|c|c|c|} \multicolumn{1}{c}{I_0} & \multicolumn{1}{c}{I-I_0} & \multicolumn{1}{c}{J_0} & \multicolumn{1}{c}{J-J_0} \\ \hline \begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix} & O & B_{11} & B_{12} \\ \hline O & \begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix} & B_{21} & B_{22} \\ \hline \end{array}}_{I \qquad\qquad\qquad J} \cdot \boldsymbol{x} = \mathbf{0} \tag{2.15.26}$$

（$I-I_0$ と J_0 の列の部分は網掛け；$I-I_0$ と J_0 の列をまとめて I'）

において，係数行列に基本行演算を施して，の部分を単位行列にすることができるためには，det  $=\pm\det B_{11}\neq 0$ でなければならないということからも導かれる．

注意 2.15.9　正則行列には少なくとも1個非零要素が含まれるから，上で $|I_0|=|J_0|=k\geqq 2$ のときには，ある $i_0\in I_0$, $j_0\in J_0$ に対して $b_{i_0j_0}\neq 0$ であり，したがって，$I''=(I-\{i_0\})\cup\{j_0\}$, $J''=(J-\{j_0\})\cup\{i_0\}$ なる基底が存在し，

$$|I-I''|=|J-J''|=1, \quad |I''-I'|=|J''-J'|=k-1$$

である．すなわち，任意の二つの基底は一つの要素を枢軸とする枢軸変換を繰り返し施すことによって，一方から他方に移ることができる．

2.15.3 LU 分解

本項では，$m\times n$ 型行列 A($m\leqq n$ とする；$m>n$ の場合も行と列の役割を入れ換えれば同様に扱える)を下三角行列 L と上三角行列 U の積 $A=LU$ に表す問題を扱う．より詳しくは，$L=[l_{ij}]$ は m 次正方行列で

$$\begin{aligned} l_{ii}&=1 \qquad (i=1,\cdots,m), \\ l_{ij}&=0 \qquad (i+1\leqq j) \end{aligned} \tag{2.15.27}$$

であるとし，$U=[u_{ij}]$ は $m\times n$ 型の行列で

$$u_{ij}=0 \qquad (i\geqq j+1) \tag{2.15.28}$$

であるとする．

注意 2.15.10 LU 分解は，次項，次々項で述べる Cholesky 分解，QR 分解とともに数値計算にはなくてはならない技法であり，各種の細かな工夫を取り入れた算法がいろいろと考案されている．

まず，補助定理を二つ用意する．(ほとんど自明であり証明は必要なかろう．)

補助定理 2.15.1 $m\times n$ 型の行列 $A=[a_{ij}]$ は，$a_{11}\neq 0$ ならば，

$$\begin{aligned} U^{(1)}=[u_{ij}^{(1)}]:\quad & u_{1j}^{(1)}=a_{1j} \quad (j=1,\cdots,n), \\ & u_{ij}^{(1)}=0 \quad (i=2,\cdots,m\,;\ j=1,\cdots,n), \\ L^{(1)}=[l_{ij}^{(1)}]:\quad & l_{11}^{(1)}=1,\ l_{i1}^{(1)}=a_{i1}/a_{11} \quad (i=2,\cdots,m), \\ & l_{ij}^{(1)}=0 \quad (j=2,\cdots,m\,;\ i=1,\cdots,m) \end{aligned} \tag{2.15.29}$$

とおくと，

$$A=L^{(1)}U^{(1)}+A^{(2)} \tag{2.15.30}$$

と表される．ここで，$A^{(2)}=[a_{ij}^{(2)}]$ は，

$$\begin{aligned} a_{1j}^{(2)}&=0 \qquad (j=1,\cdots,n), \\ a_{i1}^{(2)}&=0 \qquad (i=1,\cdots,m), \\ a_{ij}^{(2)}&=a_{ij}-a_{i1}a_{1j}/a_{11} \qquad (i=2,\cdots,m\,;\ j=2,\cdots,n) \end{aligned} \tag{2.15.31}$$

である．

$$\underset{m\times n}{\begin{pmatrix} & & \\ & a_{ij} & \\ & & \end{pmatrix}} = \begin{pmatrix} 1 & O \\ \vdots & \\ a_{i1}/a_{11} & O \\ \vdots & \end{pmatrix}\begin{pmatrix} a_{11} & \cdots\ a_{1j}\ \cdots \\ O & O \end{pmatrix} + \begin{pmatrix} 0 & O \\ O & a_{ij}-a_{i1}a_{1j}/a_{11} \end{pmatrix}$$

□

注意 2.15.11　補助定理 2.15.1 における分解(2.15.30)は一意に定まる．すなわち，$u^{(1)}_{ij}=0\,(i=2,\cdots,m\,;\,j=1,\cdots,n)$，$l^{(1)}_{11}=1$，$l^{(1)}_{ij}=0\,(j=2,\cdots,m\,;\,i=1,\cdots,m)$，$a^{(2)}_{ij}=0\,(i=1$ または $j=1)$ という条件のもとでは，$L^{(1)}$，$U^{(1)}$，$A^{(2)}$ は(2.15.29)，(2.15.31)で定められるものに限る．

補助定理 2.15.2　L の列ベクトルを $\boldsymbol{l}_{\cdot j}\ (j=1,\cdots,m)$，$U$ の行ベクトルを $\boldsymbol{u}_{i\cdot}\ (i=1,\cdots,m)$ とすると，

$$LU=\sum_{k=1}^{m}\boldsymbol{l}_{\cdot k}\boldsymbol{\cdot}\boldsymbol{u}_{k\cdot}. \tag{2.15.32}$$

である．ここで，$\boldsymbol{l}_{\cdot k}\boldsymbol{\cdot}\boldsymbol{u}_{k\cdot}$ は $m\times 1$ 型行列 $\boldsymbol{l}_{\cdot k}$ と $1\times n$ 型行列 $\boldsymbol{u}_{k\cdot}$ との行列としての積である $m\times n$ 型行列である．(注：この補助定理は行列の積の定義(2.2.2)の表現を変えただけであり，L, U が三角行列でなくても成り立つ．)　□

これらの補助定理を用いれば，次の手続きによって，A の **LU 分解**(LU-factorization)ができることがわかる(ただし，途中で現れる $a^{(k)}_{kk}$ は 0 でないと仮定する)．

[1]　$A^{(1)}:=A$，$L^{(0)}:=O_{m,m}$，$U^{(0)}:=O_{m,n}$，$k:=1$ とおき，[2]へ進む．

[2]　$u_{kj}:=0\ \ (j=1,\cdots,k-1)$，$:=a^{(k)}_{kj}\ \ (j=k,\cdots,n)$，
$l_{ik}:=0\ \ (i=1,\cdots,k-1)$，$:=1\ \ (i=k)$，
$:=a^{(k)}_{ik}/a^{(k)}_{kk}\ \ (i=k+1,\cdots,m)$，
$L^{(k)}:=L^{(k-1)}$の第 k 列(零ベクトル)を上に定めた $\boldsymbol{l}_{\cdot k}$ でおきかえたもの，
$U^{(k)}:=U^{(k-1)}$ の第 k 行(零ベクトル)を上で定めた $\boldsymbol{u}_{k\cdot}$ でおきかえたもの

により $L^{(k)}$, $U^{(k)}$ を定める．$k=m$ ならば終了．$k<m$ ならば，[3]へ進む．

[3] $a_{ij}^{(k+1)}:=0$ $(i=1,\cdots,k$ あるいは $j=1,\cdots,k)$,

$$:=a_{ij}^{(k)}-l_{ik}u_{kj}=a_{ij}^{(k)}-a_{ik}^{(k)}a_{kj}^{(k)}/a_{kk}^{(k)}$$

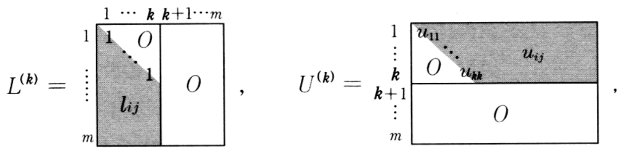
$$(i=k+1,\cdots,m\,;\ j=k+1,\cdots,n),$$

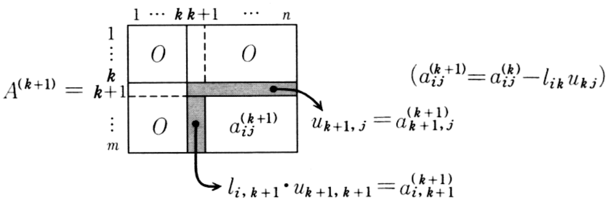
$$A^{(k+1)}:=[a_{ij}^{(k+1)}]$$

により $A^{(k+1)}$ を定め，$k:=k+1$ とおいて，[2]へ戻る．

上述の手続きの[2]において定められる l_{ij}, u_{ij} を要素とする行列，すなわち $L^{(m)}, U^{(m)}$ を L, U とすれば，(2.15.27)，(2.15.28)を満たす A の分解 $A=LU$ が得られる．図式的に表すと，$A^{(m+1)}=O_{m,n}$ と約束して，

$$A=L^{(k)}U^{(k)}+A^{(k+1)} \qquad (k=1,\cdots,m), \tag{2.15.33}$$





ということになる．

(2.15.33)から，$L, U, A^{(k)}$ の要素は A の小行列式で次のように表されることがわかる．すなわち，$|A|_{1\cdots k}^{1\cdots k}$ を作ると，三角行列の行列式(→2.6.1項)および Binet-Cauchy 展開(→2.8.3項)により，

$$|A|_{1\cdots k}^{1\cdots k}=u_{11}\cdot\cdots\cdot u_{kk} \tag{2.15.34}$$

($A^{(k+1)}$ の対応する部分行列は $O_{k,k}$ であることにも注意)であるから，

$$u_{kk}=|A|_{1\cdots k}^{1\cdots k}/|A|_{1\cdots k-1}^{1\cdots k-1} \qquad (k=1,\cdots,m) \tag{2.15.35}$$

である．($k=1$ のとき，

$$|A|_{1\cdots k-1}^{1\cdots k-1}=1$$

と定義する.)また，$j=k+1,\cdots,n$ に対して

$$|A|_{1\cdots k-1,j}^{1\cdots\cdots\cdots k}=u_{11}\cdot\cdots\cdot u_{k-1,k-1}\cdot u_{kj} \tag{2.15.36}$$

であるから，

$$u_{kj}=|A|_{1\cdots k-1,j}^{1\cdots\cdots\cdots k}/|A|_{1\cdots k-1}^{1\cdots k-1} \qquad (j\geqq k+1). \tag{2.15.37}$$

同様に，$i=k+1,\cdots,m$ に対して

$$|A|_{1\cdots\cdots\cdots k}^{1\cdots k-1,i}=l_{ik}\cdot u_{11}\cdot\cdots\cdot u_{kk} \tag{2.15.38}$$

であるから，

$$l_{ik}=|A|_{1\cdots\cdots\cdots k}^{1\cdots k-1,i}/|A|_{1\cdots k}^{1\cdots k} \qquad (i\geqq k+1) \tag{2.15.39}$$

である.

$A^{(k)}$ の要素については，(2.15.33)の両辺の部分行列の間の関係式

$$\begin{aligned} A_{1\cdots k}^{1\cdots k}&=L_{1\cdots k}^{1\cdots k}U_{1\cdots k}^{1\cdots k}, \\ A_{k+1\cdots n}^{1\cdots k}&=L_{1\cdots k}^{1\cdots k}U_{k+1\cdots n}^{1\cdots k}, \\ A_{1\cdots k}^{k+1\cdots m}&=L_{1\cdots k}^{k+1\cdots m}U_{1\cdots k}^{1\cdots k}, \\ A_{k+1\cdots n}^{k+1\cdots m}&=L_{1\cdots k}^{k+1\cdots m}U_{k+1\cdots n}^{1\cdots k}+A^{(k+1)}{}_{k+1\cdots n}^{k+1\cdots m} \end{aligned} \tag{2.15.40}$$

より，

$$\begin{aligned} L_{1\cdots k}^{k+1\cdots m}U_{k+1\cdots n}^{1\cdots k}&=A_{1\cdots k}^{k+1\cdots m}(U_{1\cdots k}^{1\cdots k})^{-1}\cdot(L_{1\cdots k}^{1\cdots k})^{-1}A_{k+1\cdots n}^{1\cdots k} \\ &=A_{1\cdots k}^{k+1\cdots m}(A_{1\cdots k}^{1\cdots k})^{-1}A_{k+1\cdots n}^{1\cdots k} \end{aligned} \tag{2.15.41}$$

に注意すれば，

$$A^{(k+1)}{}_{k+1\cdots n}^{k+1\cdots m}=A_{k+1\cdots n}^{k+1\cdots m}-A_{1\cdots k}^{k+1\cdots m}(A_{1\cdots k}^{1\cdots k})^{-1}A_{k+1\cdots n}^{1\cdots k} \tag{2.15.42}$$

を得るので，(2.10.27)(を det A で割ったもの)を参照すれば，

$$a_{ij}^{(k+1)}=|A|_{1\cdots kj}^{1\cdots ki}/|A|_{1\cdots k}^{1\cdots k} \tag{2.15.43}$$

を得る.

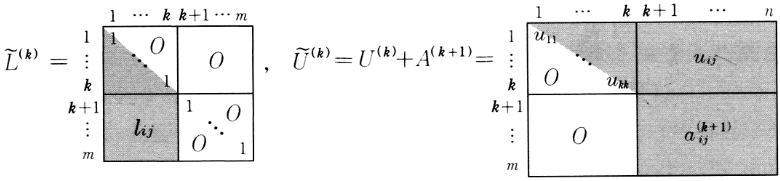
また(2.10.9)により，A が正方行列ならば，

$$\det A=|A|_{1\cdots k}^{1\cdots k}\cdot|A^{(k+1)}|_{k+1\cdots n}^{k+1\cdots n}. \tag{2.15.44}$$

この関係式は，A の代りに $A_{1\cdots k}^{1\cdots k}$ を含む A の正方部分行列に対しても，$A^{(k+1)}$ の対応する正方部分行列を取り出せば成立することは明らかであろう.

(2.15.33)，(2.15.40)は，まとめて，

$$A=\tilde{L}^{(k)}\tilde{U}^{(k)}, \tag{2.15.45}$$



とも書き表せる．ここで，$\det \tilde{L}^{(k)}=1$ であるから，$\tilde{L}^{(k)}$ は正則であり，

$$\operatorname{rank} A=k+\operatorname{rank} A^{(k+1)} \tag{2.15.46}$$

であることがわかる($u_{11},\cdots,u_{kk}\neq 0$ と仮定していることに注意)．

注意 2.15.12 以上の議論は A の LU 分解の“手続き”にはよらないことなので，分解の一意性を示したことにもなっていることに注意．

注意 2.15.13 基本変形や枢軸変換のときにも注意したように，LU 分解にも“ブロック版”があることはもちろんである．

いままでは，話を簡単にするために $u_{kk}=a_{kk}^{(k)}\neq 0$ である場合を考えてきたが，より一般の場合はどうなるであろうか．

上記の LU 分解の手続きの中で $a_{kk}^{(k)}=0$ となったとしよう．このとき，もし $a_{kj}^{(k)}\neq 0$ であるような $j\ (=k+1,\cdots,n)$ があれば，行列 A の列 k と列 j とをあらかじめ交換しておいて手続きを進めてきたと考えれば(実質的な計算はやり直す必要はない)，さらに手続きを進めることができる．すべての $j=k,\cdots,n$ に対して，$a_{kj}^{(k)}=0$ であるときには，(2.15.46)より，$\operatorname{rank} A\leqq m-1$ であることがわかる．

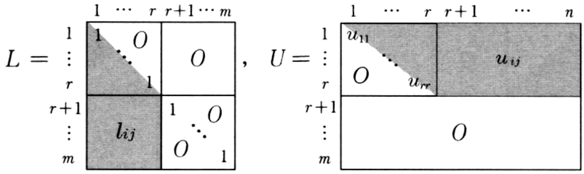
このようにして，$\operatorname{rank} A=m$ ならば，適当に列の置換をすることによって，

$$AP=LU \qquad (P \text{ は置換行列}) \tag{2.15.47}$$

という LU 分解ができる．

行の置換も許すことにすれば，ある k に対して $a_{kj}^{(k)}=0\ (j=k,\cdots,n)$ となったときでも，$a_{ij}^{(k)}\neq 0$ なる(i,j)が一つでもあるならば，行 k と i の交換，列 k と j の交換によって $a_{kk}^{(k)}\neq 0$ の場合にできる．そこで，$\operatorname{rank} A=r$ ならば，適当な m 次の置換行列 P_1，n 次の置換行列 P_2 を用いて

$$P_1AP_2=LU, \tag{2.15.48}$$



という LU 分解ができる．

注意 2.15.14　u_{ij} の代りに u_{ij}/u_{ii} を要素とする行列をあらためて U とおき，

$$D=\mathrm{diag}[u_{11},\cdots,u_{rr},1,\cdots,1] \tag{2.15.49}$$

とおくと，L も U も対角要素が 1 の三角行列（U の $r+1$ 番目以降の対角要素は 0）として

$$A=LDU,\ AP=LDU,\ P_1AP_2=LDU \tag{2.15.50}$$

という形の分解ができる．これを **LDU 分解**（LDU-factorization）と呼ぶ．

2.15.4　Cholesky 分解

a．　対称行列　　n 次対称行列 A の LU 分解を考える．前項と同様，まず簡単のため $u_{11},\cdots,u_{nn}\neq 0$ であったとしてみる．すると LDU 分解（注意 2.15.14）により

$$A=LDU \tag{2.15.51}$$

であるが，A が対称であるから

$$A=A^{\top}=U^{\top}D^{\top}L^{\top}=U^{\top}DL^{\top} \tag{2.15.52}$$

とも書ける（D は対角であるから $D^{\top}=D$）．U が上三角であるから $U^{\top}$ は下三角，同様に，$L^{\top}$ は上三角である．LDU 分解の一意性から，結局

$$L^{\top}=U,\quad U^{\top}=L \tag{2.15.53}$$

でなければならない．すなわち，n 次対称行列 A は，（要素の値の間に特殊な関係がなければ）L を対角要素が 1 に等しい下三角行列，D を対角行列として，

$$A=LDL^{\top} \tag{2.15.54}$$

という形に書き表せる．これを **Cholesky**（コレスキー，ホレスキー）**分解**（Cholesky factorization）という．

一般の場合に成り立つ Cholesky 風の分解は次の手続きによる．

［1］　$A^{(1)}:=A$，$L^{(0)}:=O_{n,n}$，$D^{(0)}:=O_{n,n}$，$k:=1$ とおき，［2］へ進む．

[2]

[2.1] $a_{ii}^{(k)} \neq 0$ となる $i\ (=k,\cdots,n)$ が存在すれば，行 i と k を入れ換え，同時に列 i と k を入れ換えた後($a_{kk}^{(k)} \neq 0$ なら入れ換え不要)，

$d_{kk} := a_{kk}^{(k)}$,

$l_{ik} := 0\ (i=1,\cdots,k-1), \quad := 1\ (i=k)$,

$\quad := a_{ik}^{(k)}/a_{kk}^{(k)}\ (i=k+1,\cdots,n)$,

$L^{(k)} := L^{(k-1)}$ の列 k を上に定めた $\boldsymbol{l}_{\cdot k}$ でおきかえたもの，

$D^{(k)} := D^{(k-1)}$ の第 k 対角要素を d_{kk} でおきかえたもの

により $L^{(k)}, D^{(k)}$ を定める．$k=n$ ならば終了；$k<n$ ならば，[3.1]へ進む．

[2.2] すべての $i\ (=k,\cdots,n)$ に対して，$a_{ii}^{(k)}=0$ であるが，$a_{ij}^{(k)}=a_{ji}^{(k)} \neq 0$ $(i \neq j\ ;\ i,j=k,\cdots,n)$ なる非対角要素があるときには，行 i,j と行 $k,k+1$ を入れ換え，同時に列 i,j と列 $k,k+1$ を入れ換えたのち(i と j のどちらを k 番目にするかはどうでもよい；$a_{k,k+1}^{(k)}=a_{k+1,k}^{(k)} \neq 0$ ならば入れ換え不要)，

$d_{kk} := d_{k+1,k+1} := 0, \quad d_{k,k+1} := d_{k+1,k} := a_{k,k+1}^{(k)} = a_{k+1,k}^{(k)}$,

$l_{ik} := 0\ (i=1,\cdots,k-1), \quad := 1\ (i=k), \quad := 0\ (i=k+1)$,

$\quad := a_{i,k+1}^{(k)}/a_{k,k+1}^{(k)}\ (i=k+2,\cdots,n)$,

$l_{i,k+1} := 0\ (i=1,\cdots,k), \quad := 1\ (i=k+1)$,

$\quad := a_{i,k}^{(k)}/a_{k,k+1}^{(k)}\ (i=k+2,\cdots,n)$,

$L^{(k+1)} := L^{(k-1)}$ の列 k および $k+1$ を上に定めた $\boldsymbol{l}_{\cdot k}$ および $\boldsymbol{l}_{\cdot k+1}$ で，それぞれ，おきかえたもの，

$D^{(k+1)} := D^{(k-1)}$ の行 $k,k+1$，列 $k,k+1$ の 2×2 対角ブロックを上記の $d_{k,k}$，$d_{k,k+1}=d_{k+1,k}$，$d_{k+1,k+1}$ でおきかえたもの

により，$L^{(k+1)}, D^{(k+1)}$ を定める．$k+1=n$ ならば終了；$k+1<n$ ならば，[3.2]へ進む．

[2.3] すべての $i,j\ (=k,\cdots,n)$ に対して $a_{ij}^{(k)}=0$ ならば終了．

[3]

[3.1]

$a_{ij}^{(k+1)} := 0\ (i=1,\cdots,k$ あるいは $j=1,\cdots,k)$,

$\quad := a_{ij}^{(k)} - l_{ik}a_{kj}^{(k)} = a_{ij}^{(k)} - a_{ik}^{(k)}a_{kj}^{(k)}/a_{kk}^{(k)}\ (i,j=k+1,\cdots,n)$,

$$A^{(k+1)} := [a_{ij}^{(k+1)}]$$

により $A^{(k+1)}$ を定め，$k := k+1$ として[2]へ戻る．

[3.2]

$$a_{ij}^{(k+2)} := 0 \ (i=1,\cdots,k+1 \text{ あるいは } j=1,\cdots,k+1),$$

$$:= a_{ij}^{(k)} - (a_{ik}^{(k)} a_{k+1,j}^{(k)} + a_{i,k+1}^{(k)} a_{k,j}^{(k)}) / a_{k,k+1}^{(k)} \ (i,j=k+2,\cdots,n),$$

$$A^{(k+2)} := [a_{ij}^{(k+2)}]$$

により $A^{(k+2)}$ を定め，$k := k+2$ として[2]へ戻る．

注意 2.15.15　A が対称なら，上記の手続きにより作られる $A^{(k)}$ も対称であることは明らかである．したがって，実際の計算は $A^{(k)}$ の対角要素と左下三角部分の要素だけを計算しておけばよい．したがって，総計算量は $n^3/6$ 回程度の乗算と加減算(除算は n 回)である．

注意 2.15.16　上の手続きにおいて，[2.1]，[3.1]の部分は，$a_{kk}^{(k)}$ を枢軸要素として，[2.2]，[3.2]の部分は，行 $k, k+1$ と列 $k, k+1$ の 2×2 対角ブロックを枢軸要素として，それぞれ，LU 分解の手続きに相当することをおこなっているにすぎないことに注意されたい．

上記の手続きが[2.1]あるいは[2.2]で終了したときには，rank $A=n$ で，分解

$$PAP^{\top} = LDL^{\top} \tag{2.15.55}$$

が達成される．ここで，L は $l_{ii}=1\ (i=1,\cdots,n)$ を満たす下三角行列，D は 1×1 型および(対称な)2×2 型対角ブロックからなるブロック対角行列(2×2 対角ブロック $\begin{bmatrix} 0 & d_{k,k+1} \\ d_{k+1,k} & 0 \end{bmatrix}$ に対応する L の対角部分は $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ である)，P は[2]においておこなった行の置換，列の置換を表す置換行列である．(なお，[2]では，行の入れ換えと同じ列の入れ換えをおこなっているので，入れ換えを表す互換行列を $P_1, P_2, \cdots$ とすると($P_i^{\top} = P_i$ である！)，行の置換は $\cdots P_2 P_1 A$ の形で，列の置換は $AP_1P_2\cdots$ の形で表される．ここで，$P=\cdots P_2P_1$ と書けば $P^{\top} = P_1^{\top} P_2^{\top} \cdots = P_1P_2\cdots$ となることにも注意．)

さらに，

$$\begin{bmatrix} 0 & d \\ d & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} d/2 & 0 \\ 0 & -d/2 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \tag{2.15.56}$$

であるから，分解(2.15.55)において，D を対角行列としておいて，L を(1×1型および 2×2 型の対角ブロックをもつ)ブロック下三角行列とすることもできる．このとき，L の 2×2 対角ブロックは $\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$ という形で，それに対応する D の二つの対角要素は絶対値が等しく符号が反対である．また，L が行 $k,k+1$，列 $k,k+1$ に対角ブロックをもつとき，それらの列の要素は上記の手続きの[2.2]で定められた $l_{ik}, l_{i,k+1}$ を用いれば，それぞれ，

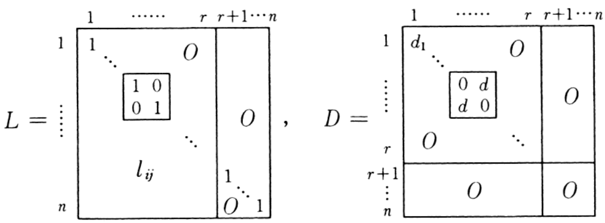
$$\begin{aligned} &l_{ik}+l_{i,k+1},\\ &l_{ik}-l_{i,k+1} \end{aligned} \tag{2.15.57}$$

と表される．

手続きが[2.3]で終了したときには，$r=\mathrm{rank}\,A=k-1$ であって，「$D:=D^{(k-1)}$，$L:=L^{(k-1)}$ の列 $k,\cdots,n$ を $l_{ii}:=1$ $(i=k,\cdots,n)$，$l_{ij}:=0$ $(j=k,\cdots,n;\ i\neq j,\ i=1,\cdots,n)$ でおきかえたもの」とすれば，(2.15.55)のような分解が成り立つ．

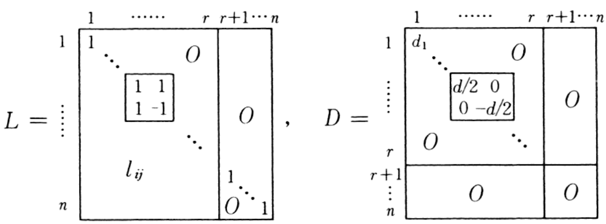
以上を図式的に描けば，

$$PAP^{\mathsf{T}}=LDL^{\mathsf{T}},$$



(2.15.58)

あるいは



(2.15.59)

である．

b．Hermite 行列　$K=\mathbb{C}$ で A が Hermite 行列($A^{\mathsf{H}}=A$；対角要素 a_{ii} はすべて実数，非対角要素は $a_{ji}=\overline{a_{ij}}$ $(i>j)$ を満たす)のときにも，A が対称のと

きとほとんど同じ Cholesky 風の分解 $A=LDL^{\mathrm{H}}$ ができることは明らかであろう．(途中で現れる $A^{(k)}$ が Hermite 行列になること，実質的計算を対角要素と左下三角部分についてだけおこなえばよいことも同様.) Cholesky 分解の手続きの中で，[2.2]と[3.2]のところが次のように変更される．

[2.2]′　すべての $i\ (=k,\cdots,n)$ に対して，$a_{ii}^{(k)}=0$ であるが，$\overline{a_{ij}^{(k)}}=a_{ji}^{(k)}\neq 0$ $(i\neq j\ ;\ i,j=k,\cdots,n)$ なる非対角要素があるときには，行 i,j と行 $k,k+1$ とを入れ換え，同時に列 i,j と列 $k,k+1$ とを入れ換えたのち (i と j のどちらを k 番目にするかはどうでもよい；$\overline{a_{k,k+1}^{(k)}}=a_{k+1,k}^{(k)}\neq 0$ ならば入れ換え不要)，

$d_{kk}:=d_{k+1,k+1}:=0,\quad d_{k+1,k}:=a_{k+1,k}^{(k)},\quad d_{k,k+1}:=\overline{d_{k+1,k}},$

$l_{ik}:=0\ (i=1,\cdots,k-1),\quad :=1\ (i=k),\quad :=0\ (i=k+1),$

$\quad :=a_{i,k+1}^{(k)}/a_{k,k+1}^{(k)}\ (i=k+2,\cdots,n),$

$l_{i,k+1}:=0\ (i=1,\cdots,k),\quad :=1\ (i=k+1),$

$\quad :=a_{i,k}^{(k)}/a_{k+1,k}^{(k)}\ (i=k+2,\cdots,n),$

$L^{(k+1)}:=L^{(k-1)}$ の列 k および $k+1$ を上に定めた $\boldsymbol{l}_{\cdot k}$ および $\boldsymbol{l}_{\cdot k+1}$ で，それぞれ，おきかえたもの，

$D^{(k+1)}:=D^{(k-1)}$ の行 $k,k+1$，列 $k,k+1$ の 2×2 対角ブロックを上記の $d_{k,k},d_{k,k+1},d_{k+1,k},d_{k+1,k+1}$ でおきかえたもの

により，$L^{(k+1)},D^{(k+1)}$ を定める．$k+1=n$ ならば終了；$k+1<n$ ならば，[3.2]′ へ進む．

[3.2]′

$a_{ij}^{(k+2)}:=0\ (i=1,\cdots,k+1$ あるいは $j=1,\cdots,k+1),$

$\quad :=a_{ij}^{(k)}-a_{ik}^{(k)}\cdot a_{k+1,j}^{(k)}/a_{k+1,k}^{(k)}-a_{i,k+1}^{(k)}\cdot a_{kj}^{(k)}/a_{k,k+1}^{(k)}\ (i,j=k+2,\cdots,n),$

$A^{(k+2)}:=[a_{ij}^{(k+2)}]$

により $A^{(k+2)}$ を定め，$k:=k+2$ として，[2]へ戻る．

さらに，$d=d_{k+1,k}\ (\bar{d}=d_{k,k+1})$ が $d=\rho e^{i\theta}\ (\bar{d}=\rho e^{-i\theta})$ であるとき ($\rho>0$，θ は

実数),

$$\begin{bmatrix} 0 & \bar{d} \\ d & 0 \end{bmatrix} = \begin{bmatrix} 1 & \mathrm{e}^{-\mathrm{i}\theta} \\ \mathrm{e}^{\mathrm{i}\theta} & -1 \end{bmatrix} \begin{bmatrix} \rho/2 & 0 \\ 0 & -\rho/2 \end{bmatrix} \begin{bmatrix} 1 & \mathrm{e}^{-\mathrm{i}\theta} \\ \mathrm{e}^{\mathrm{i}\theta} & -1 \end{bmatrix} \tag{2.15.60}$$

であるから，分解 $A=LDL^{\mathrm{H}}$ において，D の方を対角行列として，L をブロック下三角行列とすることもできる．このときには，L の 2×2 対角ブロックは $\begin{bmatrix} 1 & \mathrm{e}^{-\mathrm{i}\theta} \\ \mathrm{e}^{\mathrm{i}\theta} & -1 \end{bmatrix}$ という形で，それに対応する D の二つの対角要素は絶対値の等しい正の数と負の数である．また，L が行 $k, k+1$，列 $k, k+1$ に対角ブロックをもつとき，それらの列の要素は上記の[2.2]′ で定められた $l_{ik}, l_{i,k+1}$ を用いれば，それぞれ，

$$\begin{aligned} &l_{ik} + \mathrm{e}^{\mathrm{i}\theta} l_{i,k+1}, \\ &\mathrm{e}^{-\mathrm{i}\theta} l_{ik} - l_{i,k+1} \end{aligned} \tag{2.15.61}$$

と表される(ここで $\theta = \arg a_{k+1,k}^{(k)}$)．

分解手続きの途中に現れる行列 $A^{(k)}$ が Hermite 行列であることも明らかであろう．

注意 2.15.17 対称行列あるいは Hermite 行列 A の階数が r であるときには，その r 次の主小行列式 $|A|_{i_1\cdots i_r}^{i_1\cdots i_r}$ の中に 0 でないものが存在する．(階数の定義から，$r+1$ 次以上のものは 0 である．また，$r-1$ 次以下の小行列式の中に 0 でないものが存在するが，それが主小行列式であるかどうかはわからない．たとえば，$A=\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ の階数は 2 で $\det A = |A|_{12}^{12} = -1 \neq 0$ であるが，1 次の主小行列式は $|A|_1^1 = 0$，$|A|_2^2 = 0$ である．) このことも分解(2.15.58)，(2.15.59) あるいはその Hermite 版から示すことができる．すなわち，置換 P により $1, \cdots, r$ となる行，列の番号を $i_1, \cdots, i_r$ とし，$|A|_{i_1\cdots i_r}^{i_1\cdots i_r}$ に Binet-Cauchy の公式(2.8.17)を適用すると，D の r 次の小行列式の中で 0 にならないものが $|D|_{1\cdots r}^{1\cdots r}$ だけであることに注意すれば，

$$|A|_{i_1\cdots i_r}^{i_1\cdots i_r} = |L|_{1\cdots r}^{1\cdots r} \, |D|_{1\cdots r}^{1\cdots r} \, |L^{\mathrm{T}}|_{1\cdots r}^{1\cdots r} \tag{2.15.62}$$

であり，$|L|_{1\cdots r}^{1\cdots r} = |L^{\mathrm{T}}|_{1\cdots r}^{1\cdots r} \neq 0$ も明らかであるから，$|A|_{i_1\cdots i_r}^{i_1\cdots i_r} \neq 0$ である．(A が Hermite 行列のときには L^{T} の代りに L^{H} を用いる．)

c. 反対称行列　n 次反対称行列 A に対しても，対称行列に対する分解とほとんど同様の分解が可能であることは容易に想像されよう．反対称行列の対角要素はすべて 0 であるから，本項 a における分解手続きのうち，[2.1]，[3.1] に相当する場合は生じない．そこで次の通りの手続きが得られる．

[1]　$A^{(1)} := A$, $L^{(0)} := O_{n,n}$, $D^{(0)} := O_{n,n}$, $k := 1$ とおき，[2]へ進む．

[2]　$A^{(k)} = O$ ならば終了．$a_{ij}^{(k)} = -a_{ji}^{(k)} \neq 0$ $(i \neq j;\ i, j = k, \cdots, n)$ があるときには，行 i, j と行 $k, k+1$ を入れ換え，同時に列 i, j と列 $k, k+1$ を入れ換えたのち(i と j のどちらを k 番目にするかはどうでもよい；$a_{k,k+1}^{(k)} = -a_{k+1,k}^{(k)} \neq 0$ なら入れ換え不要)，

$d_{kk} := d_{k+1,k+1} := 0$,　$d_{k+1,k} := a_{k+1,k}^{(k)}$,　$d_{k,k+1} := -d_{k+1,k}$,

$l_{ik} := 0$ $(i = 1, \cdots, k-1)$,　$:= 1$ $(i = k)$,　$:= 0$ $(i = k+1)$,

$:= a_{i,k+1}^{(k)} / a_{k,k+1}^{(k)}$ $(i = k+2, \cdots, n)$,

$l_{i,k+1} := 0$ $(i = 1, \cdots, k)$,　$:= 1$ $(i = k+1)$,

$:= a_{i,k}^{(k)} / a_{k+1,k}^{(k)}$ $(i = k+2, \cdots, n)$,

$L^{(k+2)} := L^{(k-1)}$ の列 k および $k+1$ を上に定めた $\boldsymbol{l}_{\cdot k}$ および $\boldsymbol{l}_{\cdot k+1}$ で，それぞれ，おきかえたもの，

$D^{(k+1)} := D^{(k-1)}$ の第 $k, k+1$ 行と第 $k, k+1$ 列の 2×2 対角ブロックを上記の d_{ij} $(i, j = k, k+1)$ でおきかえたもの

により $L^{(k+1)}$，$D^{(k+1)}$ を定める．$k+1 = n$ ならば終了；$k+1 < n$ ならば，[3]へ進む．

[3]

$a_{ij}^{(k+2)} := 0$ $(i = 1, \cdots, k+1$ あるいは $j = 1, \cdots, k+1)$,

$:= a_{ij}^{(k)} - (a_{ik}^{(k)} \cdot a_{k+1,j}^{(k)} - a_{i,k+1}^{(k)} \cdot a_{kj}^{(k)}) / a_{k+1,k}^{(k)}$ $(i, j = k+2, \cdots, n)$,

$A^{(k+2)} := [a_{ij}^{(k+2)}]$

により $A^{(k+2)}$ を定め，$k := k+2$ として，[2]へ戻る．

さらに，

$$\begin{bmatrix} 0 & -d \\ d & 0 \end{bmatrix} = \begin{bmatrix} d & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} d & 0 \\ 0 & 1 \end{bmatrix} \qquad (2.15.63)$$

であるから，[2]において，

$$d_{k+1,k} := 1, \quad d_{k,k+1} := -1, \quad d_{k,k} := d_{k+1,k+1} := 0$$

とおき，l_{ik} を $a_{k+1,k}^{(k)}$ 倍したものをあらためて l_{ik} とすることもできる($l_{i,k+1}$ はそのまま)．

図式的に描けば(P は置換行列)，

$$PAP^{\top}=LDL^{\top},$$

$$L=\begin{array}{c} \\ 1\\ 2\\ \vdots\\ \\ n \end{array}\!\!\begin{array}{c} \begin{array}{ccccc} 1\ \ 2 & \cdots\cdots & r & r+1\cdots n \end{array}\\ \left(\begin{array}{ccc|c} \begin{smallmatrix}1&0\\0&1\end{smallmatrix} & & O & \\ & \begin{smallmatrix}1&0\\0&1\end{smallmatrix} & & O\\ & & \ddots & \\ & l_{ij} & & \begin{smallmatrix}1&&\\&\ddots&\\O&&1\end{smallmatrix} \end{array}\right)\end{array},\qquad D=\begin{array}{c} \\ 1\\ 2\\ \vdots\\ r\\ r+1\\ \vdots\\ n \end{array}\!\!\begin{array}{c} \begin{array}{ccccc} 1\ \ 2 & \cdots\cdots & r & r+1\cdots n \end{array}\\ \left(\begin{array}{ccc|c} \begin{smallmatrix}0&-d_1\\d_1&0\end{smallmatrix} & & O & \\ & \begin{smallmatrix}0&-d_2\\d_2&0\end{smallmatrix} & & O\\ O & & \ddots & \\ \hline & O & & O \end{array}\right)\end{array},\tag{2.15.64}$$

あるいは

$$L=\begin{array}{c} \\ 1\\ 2\\ \vdots\\ \\ n \end{array}\!\!\begin{array}{c} \begin{array}{ccccc} 1\ \ 2 & \cdots\cdots & r & r+1\cdots n \end{array}\\ \left(\begin{array}{ccc|c} \begin{smallmatrix}d_1&0\\0&1\end{smallmatrix} & & O & \\ & \begin{smallmatrix}d_2&0\\0&1\end{smallmatrix} & & O\\ & & \ddots & \\ & l_{ij} & & \begin{smallmatrix}1&&\\&\ddots&\\O&&1\end{smallmatrix} \end{array}\right)\end{array},\qquad D=\begin{array}{c} \\ 1\\ 2\\ \vdots\\ r\\ r+1\\ \vdots\\ n \end{array}\!\!\begin{array}{c} \begin{array}{ccccc} 1\ \ 2 & \cdots\cdots & r & r+1\cdots n \end{array}\\ \left(\begin{array}{ccc|c} \begin{smallmatrix}0&-1\\1&0\end{smallmatrix} & & O & \\ & \begin{smallmatrix}0&-1\\1&0\end{smallmatrix} & & O\\ O & & \ddots & \\ \hline & O & & O \end{array}\right)\end{array}\tag{2.15.65}$$

である．

注意 2.15.18 対称行列の Cholesky 分解(本項 a)と Hermite 行列の分解(本項 b)との関係に類似のことは，反対称行列と歪 Hermite 行列との間については成り立たない．A が歪 Hermite なら iA が Hermite であるから，iA の変形を考えればよい．ただ，(2.15.64)あるいは(2.15.65)の分解を遂行するために要する計算量がほぼ $n^3/6$ 程度であることは注意 2.15.15 に述べたのと同様である．

注意 2.15.19 以上の具体的な分解手続きに基づくと，2.9 節で述べた反対称行列の階数や Pfaffian についての話が見通しよく理解できる．

すなわち，行列の要素が属する体 K の標数がたとえ 2 であっても，反対称行列の階数 rank A は，L が正則であるから，rank $D=r$ に等しく，r は分解の手続きからも明らかなように**偶数**である．したがって，奇数次の反対称行列の行列式は 0 である．

$r=n=2m$ のときの det A の表式を調べよう．(2.15.65)において $A=(P^{-1}L)D(P^{-1}L)^{\top}$ と表して(A，$P^{-1}L$ が注意 2.9.3 の B，A に対応する！)，(2.9.16)を適用すれば，$2m$ 次の反対称行列 A に対しては Pf $A=\det\,(P^{-1}L)$ であり，det D

$=1$ であるから $\det A=(\det(P^{-1}L))^2\cdot(\det D)=(\mathrm{Pf}\,A)^2$ である．

反対称行列 A の階数が $n=2m$ でなくて $r=2q$ $(<n)$ のときに，分解(2.15.65)により，$\tilde{A}=PAP^{\mathrm{T}}$ の要素は

$$\tilde{a}_{\tilde{i}\tilde{j}}=\sum_{k=1}^{q}(l_{\tilde{i}}^{2k}l_{\tilde{j}}^{2k-1}-l_{\tilde{i}}^{2k-1}l_{\tilde{j}}^{2k})=2\sum_{k=1}^{q}l_{[\tilde{i}}^{2k}l_{\tilde{j}]}^{2k-1} \tag{2.15.66}$$

と表せる．そこで注意 2.9.1 で定義した**小 Pfaffian**

$$(\mathrm{Pf}\,A)_{i_1\cdots i_{2q}}=\frac{(2q)!}{q!2^q}a_{[i_1 i_2}a_{i_3 i_4}\cdots a_{i_{2q-1}i_{2q}]} \tag{2.15.67}$$

を計算してみる．(i,j と $\tilde{i},\tilde{j}$ とは置換 P で対応づけられているものとする．その対応のもとで，$i_1,\cdots,i_{2q}$ と $\tilde{i}_1,\cdots,\tilde{i}_{2q}$ との並び順を揃えるのに必要な置換の偶奇の符号を $\varepsilon_{\tilde{i}_1\cdots\tilde{i}_{2q}}^{i_1\cdots i_{2q}}$ で表すことにする．)

$$(\mathrm{Pf}\,\tilde{A})_{\tilde{i}_1\cdots\tilde{i}_{2q}}=\varepsilon_{\tilde{i}_1\cdots\tilde{i}_{2q}}^{i_1\cdots i_{2q}}(\mathrm{Pf}\,A)_{i_1\cdots i_{2q}} \tag{2.15.68}$$

であるが，(2.15.66)によれば，これはさらに

$$=\frac{(2q)!}{q!}\left(\sum_{k_1=1}^{q}l_{[\tilde{i}_1}^{2k_1}l_{\tilde{i}_2}^{2k_1-1}\right)\cdots\left(\sum_{k_q=1}^{q}l_{\tilde{i}_{2q-1}}^{2k_q}l_{\tilde{i}_{2q}]}^{2k_q-1}\right) \tag{2.15.69}$$

と表せる．(2.9.16)を導いたのと同様な考察により，この括弧をほどいて得られる各項

$$(2q)!\,l_{[\tilde{i}_1}^{2k_1}l_{\tilde{i}_2}^{2k_1-1}\cdots l_{\tilde{i}_{2q-1}}^{2k_q}l_{\tilde{i}_{2q}]}^{2k_q-1} \tag{2.15.70}$$

は，$k_1,\cdots,k_q$ の中に等しいものがあるときには 0 となり，それらがすべて異なるとき，すなわち $\{k_1,\cdots,k_q\}=\{1,\cdots,q\}$ であるときには

$$(2q)!\,l_{[\tilde{i}_1}^{1}l_{\tilde{i}_2}^{2}\cdots l_{\tilde{i}_{2q-1}}^{2q-1}l_{\tilde{i}_{2q}]}^{2q} \tag{2.15.71}$$

の $(-1)^q$ 倍になる．このような項は $(1,\cdots,q)$ のすべての順列について加え合わされるから，さらに $q!$ 倍され結局

$$\begin{aligned}(\mathrm{Pf}\,\tilde{A})_{\tilde{i}_1\cdots\tilde{i}_{2q}}&=\pm(-1)^q(2q)!\,l_{[i_1}^{1}l_{i_2}^{2}\cdots l_{i_{2q}]}^{2q}\\&=\varepsilon_{\tilde{i}_1\cdots\tilde{i}_{2q}}^{i_1\cdots i_{2q}}|L|_{1,\cdots,2q}^{i_1\cdots i_{2q}}\end{aligned} \tag{2.15.72}$$

を得る．ところで，$|D|_{k_1\cdots k_{2q}}^{k_1\cdots k_{2q}}$ の中で 0 にならないものは $|D|_{1,\cdots,2q}^{1,\cdots,2q}=1$ のみであるから(これが $|L|_{1,\cdots,2q}^{\cdots\cdots}$ のみを考えた理由である)，Binet-Cauchy の公式と(2.15.65)とにより，

$$|\tilde{A}|_{\tilde{j}_1\cdots\tilde{j}_{2q}}^{\tilde{i}_1\cdots\tilde{i}_{2q}}=\varepsilon_{\tilde{i}_1\cdots\tilde{i}_{2q}}^{i_1\cdots i_{2q}}\varepsilon_{\tilde{j}_1\cdots\tilde{j}_{2q}}^{j_1\cdots j_{2q}}|L|_{1,\cdots,2q}^{i_1\cdots i_{2q}}|L|_{1,\cdots,2q}^{j_1\cdots j_{2q}}. \tag{2.15.73}$$

よって

$$\begin{aligned}|A|_{j_1\cdots j_{2q}}^{i_1\cdots i_{2q}}&=\varepsilon_{\tilde{i}_1\cdots\tilde{i}_{2q}}^{i_1\cdots i_{2q}}\varepsilon_{\tilde{j}_1\cdots\tilde{j}_{2q}}^{j_1\cdots j_{2q}}|\tilde{A}|_{\tilde{j}_1\cdots\tilde{j}_{2q}}^{\tilde{i}_1\cdots\tilde{i}_{2q}}\\&=(\mathrm{Pf}\,A)_{i_1\cdots i_{2q}}(\mathrm{Pf}\,A)_{j_1\cdots j_{2q}}\end{aligned} \tag{2.15.74}$$

という関係式が得られる．

2.15.5 QR 分 解

実数体 $\mathbb{R}$ の上の $m\times n$ 型の行列 A は，m 次直交行列 Q と $m\times n$ 型の上三角

行列 $R\,(r_{ij}=0, i>j)$ との積

$$A=QR \tag{2.15.75}$$

に分解できる．これを QR 分解(QR factorization)と呼ぶ．QR 分解にはいくつかのやり方があるが，ここでは，例 2.14.4 の Householder 行列を利用する方法を述べる．A が複素数体の上の行列で Q がユニタリ行列とした場合も，例 2.14.6 の複素版の Householder 行列を利用すればまったく同じように話を進めることができるので，ここでは簡単のために，A が実数行列で Q が直交行列の場合だけを扱う．

考え方は簡単で，まず A の第 1 列の列ベクトルを，それと長さが等しくて第 2 行以下の要素が 0 であるようなベクトルに移す Householder 行列 $Q^{(1)}$ を作り，$A'=Q^{(1)}A$ とする；次に，A' の第 2 行以下だけに注目し，その第 2 列の列ベクトルの第 3 行以下を 0 にするような $Q^{(2)}$ を作り，$A''=Q^{(2)}A'$ とする；…；という方針で臨む．詳しくは次の手続きによる．

[1] $A^{(1)}:=A$,　$k:=1$ とおき，[2]へ進む．

[2]

[2.1] $A^{(k)}=O$ ならば，[4]へ進む．

[2.2] $A^{(k)}\neq O$ ならば，$a_{ij}^{(k)}\neq 0$ なる $i(\in\{k,\cdots,m\})$ が存在するような列 j を選び，それと列 k とを入れ換える($a_{ik}^{(k)}\neq 0$ なる $i(\in\{k,\cdots,m\})$ が存在すれば入れ換え不要)．$k=m$ ならば，[4] へ，$k<m$ ならば，[3] へ進む．

[3]

$$\sigma^{(k)}:=\left(\sum_{i=k}^{m}|a_{ik}^{(k)}|^2\right)^{1/2}, \qquad \boldsymbol{b}^{(k)}:=[\underset{k}{\pm\sigma^{(k)}}, \underset{k+1}{0}, \underset{\cdots}{\cdots}, \underset{m}{0}]^{\top} \tag{2.15.76}$$

により列ベクトル $\boldsymbol{b}^{(k)}$ を定める．(要素番号を $k, k+1, \cdots, m$ としていることに注意；$\pm$ は次式(2.15.77)の第 1 式の計算で桁落ちが生じないように，すなわち $a_{kk}^{(k)}$ の符号と反対に選ぶ．)

$$
\begin{aligned}
\boldsymbol{v}^{(k)} &:= \boldsymbol{b}^{(k)} - \boldsymbol{a}_{\cdot k}^{(k)}, \\
\boldsymbol{u}^{(k)} &:= \boldsymbol{v}^{(k)} \bigg/ \left(\sum_{i=k}^{m} |v_i^{(k)}|^2 \right)^{1/2}
\end{aligned}
\tag{2.15.77}
$$

により $\boldsymbol{u}^{(k)}$ を定め，$H(\boldsymbol{u}^{(k)})A^{(k)}$ $(H(\boldsymbol{u}^{(k)}) = I_{m-k+1} - 2\boldsymbol{u}^{(k)}\boldsymbol{u}^{(k)\mathrm{T}})$ の最上行と最左列を除いた$(m-k)\times(n-k)$型の行列を $A^{(k+1)} = [a_{ij}^{(k+1)}]$ $(i=k+1,\cdots,m\,;\ j=k+1,\cdots,n)$ とする．

$k:=k+1$ とし，[2]へ戻る．

[4] $h:=1,\cdots,k-1$ に対し，

$$
Q^{(h)} = \begin{array}{c}
 \\ 1 \\ \vdots \\ h-1 \\ h \\ \vdots \\ m
\end{array}
\begin{array}{c}
\begin{array}{cc} 1\ \cdots\ h-1 & h\ \cdots\cdots\ m \end{array} \\
\left[\begin{array}{c|c}
\begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix} & O \\ \hline
O & H(\boldsymbol{u}^{(h)})
\end{array}\right]
\end{array}
\tag{2.15.78}
$$

により $Q^{(h)}$ (明らかに直交行列)を定め，

$$
Q = Q^{(1)} \cdots Q^{(k-1)} \tag{2.15.79}
$$

(これも直交行列)とおく．また，$h:=1,\cdots,k-1$ に対しては

$$
\begin{aligned}
&r_{hj} := 0 \qquad (j=1,\cdots,h-1), \\
&[r_{hh},\cdots,r_{hn}] := H(\boldsymbol{u}^{(h)})A^{(h)} \text{ の最上行},
\end{aligned}
\tag{2.15.80}
$$

$h:=k$ に対しては

$$
\begin{aligned}
&r_{kj} := 0 \qquad (j=1,\cdots,k-1), \\
&[r_{kk},\cdots,r_{kn}] := A^{(k)} \text{ の最上行},
\end{aligned}
\tag{2.15.81}
$$

$h:=k+1,\cdots,m$ に対しては

$$
r_{hj} := 0 \qquad (j=1,\cdots,n) \tag{2.15.82}
$$

により $R=[r_{ij}]$ を定める．

上の手続きにより，[2.1]でおこなった列の入れ換えを表す n 次置換行列を P とすれば，

$$
AP = QR \tag{2.15.83}
$$

という分解が得られたことになる．(実際には，A に $Q^{(1)}, Q^{(2)}, \cdots, Q^{(k-1)}$ をこの

順に左から掛けた $Q^{(k-1)}\cdots Q^{(1)}A$ が R になるわけであるが，$(Q^{(h)})^{\top}=Q^{(h)}$ であるから $(Q^{(k-1)}\cdots Q^{(1)})^{\top}=Q$ となる．）

注意 2.15.20 上の手続きから明らかなように，

$$\mathrm{rank}\,A=\mathrm{rank}\,R \tag{2.15.84}$$

である．また，A の各列の“長さ”と R の対応する列の長さは等しい．さらに，

$$r_{hh}=\pm\sigma^{(h)} \qquad (h=1,\cdots,k). \tag{2.15.85}$$

上の手続きに必要とされる演算の回数は，m,k が共に n に等しいとき，約 $2n^3/3$ 回の加減算と同数の乗算，それに約 n 回の除算と開平とである．

A の列の置換 P を固定したとき，$\mathrm{rank}\,A=m$ ならば，QR 分解はほとんど一意的である．すなわち，$QR=Q'R'$，$\mathrm{rank}\,R=\mathrm{rank}\,R'=m$ ならば，R と R' の左部の m 次正方上三角部分行列を $\tilde{R},\tilde{R}'$ とすると，$Q\tilde{R}=Q'\tilde{R}'$，$Q'^{\top}Q=\tilde{R}'\tilde{R}^{-1}$ となるが，この $\tilde{R}'\tilde{R}^{-1}$ は上三角行列，したがって $Q'^{\top}Q$ は上三角直交行列である．例 2.14.1 の中で注意したように，そのような $Q'^{\top}Q$ は(2.14.12)の形の行列(Q_0 とする)である．ゆえに $Q=Q'Q_0$，すなわち，Q と Q' は各列の符号を除いて一致する．R と R' も各行の符号を除いて一致する．

実際の計算には，分解の情報は $\boldsymbol{u}^{(1)},\cdots,\boldsymbol{u}^{(m)}$ で表し行列の形で表すのではないとか，(2.15.80)などの計算も $\boldsymbol{u}^{(h)}$ を用いておこなうとかいうような，各種の注意が必要になるが，ここでは省く．

2.15.6 単 因 子

線形代数では体 K の上の行列を主として取り扱うのであるが，K の代りにある可換環，特に“Euclid(ユークリッド；ギリシア名は Eukleidēs エウクレイデース)環”(Euclid 互除法が成り立つような環)の元を要素とする行列についての知識も応用上大切なことがある．以下では，そのような行列の性質を見やすくするための標準的な変形について述べる．行列の要素は整数($\in\mathbb{Z}$)であるとして論じるが，それが体 K の上の 1 変数多項式($\in K[x]$)であっても

整数	$\longleftrightarrow$	多項式
±1($\mathbb{Z}$ の単元)	$\longleftrightarrow$	零でない K の元($K[x]$の単元)
絶対値	$\longleftrightarrow$	多項式の次数

という対応関係のもとで，すべての議論は，1 変数多項式環 $K[x]$ に対しても成り立つことに注意してほしい．

特に，最も基本的な“Euclid 互除法(Euclidean algorithm)”の一段は，$\mathbb{Z}$ に

おいては，

「任意の整数 $a, b\,(\neq 0)$ に対して，

$$a = qb + r, \quad |r| < |b|, \quad r \geqq 0 \tag{2.15.86}$$

を満たす整数 q（**商**(quotient)）と r（**剰余**(remainder)）が存在する」

であり，$K[x]$ においては，

「任意の多項式 $a(x), b(x)\,(\neq 0)$ に対して，

$$a(x) = q(x)b(x) + r(x),$$

$$\text{ここで，“}r(x)\text{の次数}<b(x)\text{の次数”} \quad \text{あるいは} \quad r(x)=0 \tag{2.15.87}$$

を満たす商多項式 $q(x)$ と剰余多項式 $r(x)$ が存在する」

である．

(2.15.86)，(2.15.87) において，$r=0$ のとき，“b は a を割り切る (b divides a)” といい，“$b|a$” と記す．また，$a\,(\neq 0), b\,(\neq 0)$ に対して，$c|a$，$c|b$ で絶対値最大（次数最大）の c のことを a と b の**最大公約数（最大公約多項式）**(greatest common divisor) と呼び，$\gcd(a,b)$ と記す．最大公約数は符号を除いて（最大公約多項式は非零定数倍を除いて）一意に定まる．（便宜上，$\mathbb{Z}$ においては最大公約数としては正の数を，$K[x]$ においては x の最高次数のべきの係数が 1 の多項式（モニック多項式，monic polynomial という）を採るのが普通である．）3 個以上の整数の最大公約数 $\gcd(a,b,c,\cdots)$ についても同様．

基本行演算（あるいは列演算）(2.15.1 項の (i)，(ii)，(iii)) の代りに，以下の**単模基本変形**(unimodular elementary transformation) を考える．

(i′)　行（列）i_0 と j_0 を交換する．

(ii′)　行（列）i_0 に -1 を掛ける（符号を変える）．

(iii′)　行 j_0（列 i_0）を t 倍して行 i_0（列 j_0）に加える．

注意 2.15.21　単模基本変形を表す行列は 2.3.2 項の記号を用いれば，$P(i_0, j_0)$, $T(i_0\,;\,-1)$ ($K[x]$ においては $T(i_0\,;\,t)$ $(t \in K-\{0\})$)，$W(i_0, j_0\,;\,t)$ (t：整数（多項式））である．$\det P(i_0, j_0) = -1$，$\det T(i_0\,;\,-1) = -1$ ($\det T(i_0\,;\,t) = t \in K-\{0\}$)，$\det W(i_0, j_0\,;\,t) = 1$ であるから，これら行列の積の行列式は 1 あるいは -1 ($K[x]$ では単元すなわち非零定数）である．これが“単模(unimodular)”の言葉の由来である（模は modulus の音に擬したもの；modulus はこの場合“絶対値”を意味する；模は規模の模である）．一般に行列式の絶対値が 1（行列式が非零

定数)であるような整数(多項式)を要素とする行列を**単模行列**(unimodular matrix)と呼ぶ．単模行列の逆行列も整数(多項式)を要素とする単模行列であることは明らかである．

2.15.1 項の基本変形や 2.15.3 項の LU 分解に類似の変形が単模基本行演算を用いて可能である．

まず，$m\times n$ 型の行列 A(要素は整数)に対して，適当な単模行列 S(m 次正方)と置換行列 P(n 次正方)とを選んで

$$SAP=U \tag{2.15.88}$$

とすることができる．ここで，$U=[u_{ij}]\,(i=1,\cdots,m\,;\,j,\cdots,n)$ は，$\mathrm{rank}\,A=r$ のとき，

$$\begin{aligned} &u_{ij}\geqq 0 && (i=1,\cdots,m\,;\,j=1,\cdots,r),\\ &u_{ij}=0 && (i>j \text{ あるいは } i>r \text{ のとき}),\\ &|u_{ij}|<|u_{jj}| && (i<j, j=1,\cdots,r) \end{aligned} \tag{2.15.89}$$

を満たす上三角行列である．

この変形のための算法を記述する前に，m 次非零ベクトル $\boldsymbol{b}=[b_1,\cdots,b_m]^{\mathrm{T}}$ は，単模基本行演算によって，$\boldsymbol{c}=[c_1,0,\cdots,0]^{\mathrm{T}}$，$c_1=\gcd(b_1,\cdots,b_m)\,(>0)$ という形にすることができることに注意しよう．それには次の GCD 算法を用いればよい．(計算複雑度の観点からこれが最良の算法というわけではない．)

GCD 算法

[1] $b_1,\cdots,b_m$ の符号をみな非負(多項式の場合はモニック)に変える．$m=1$ ならば終了；$m\geqq 2$ のとき，[2]に進む．

[2] $b_1,\cdots,b_m$ の中の非零で最小なものと b_1 とを入れ換える．

[3]

[3.1] $b_2=\cdots=b_m=0$ ならば終了．

[3.2] $b_k\neq 0$ なる $k(=2,\cdots,m)$ があれば，そのようなすべての k に対し，b_k を b_1 で割って商 q_k と剰余 $r_k(\geqq 0)$ を求め，b_1 を q_k 倍して b_k から引き(その結果 $b_k:=r_k$ となる)，[2]へ戻る(多項式の場合は b_k をモニックにしてから)．

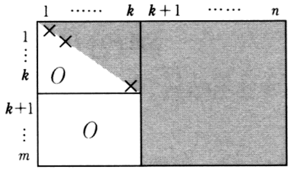
この算法において，

(イ)どの段においても，$b_1,\cdots,b_m$ の最大公約数は変化しない，

(ロ)[2]と[3.2]を通るごとに，非零の $b_2,\cdots,b_m$ の最大値(次数)が“減少”することは明らかである．(ロ)によって，この算法が有限回のステップで終了すること，(イ)によって，そのときの $\boldsymbol{b}$ が上記の $\boldsymbol{c}$ の性質をもつことがわかる．

(2.15.88)，(2.15.89)の変形をおこなうには，次のようにすればよい．ここで，$A^{(k)}$ は，次の図式のような形をしている．

$A^{(k)} =$



(2.15.90)

[1]　$A^{(0)}:=A$，$k:=0$ とおき，[2]へ進む．

[2]　$k=\min(m,n)$ ならば $r:=k$ として終了．そうでなければ

[2.1]　$a_{ij}^{(k)}=0$ $(i=k+1,\cdots,m\,;\,j=k+1,\cdots,n)$ ならば，$r:=k$ とおいて終了．

[2.2]　$a_{ij}^{(k)}\neq 0$ であるような $i(=k+1,\cdots,m)$，$j(=k+1,\cdots,n)$ が存在すれば，$A^{(k)}$ の列 $k+1$ と列 j とを入れ換え，列 $k+1$ の行 $k+1,\cdots,m$ の部分のベクトルに対して GCD 算法を適用し，同時に同じ単模基本行演算を他の列に対しても適用する(実質的には列 $k+1,\cdots,n$，行 $k+1,\cdots,m$ の部分だけが変化する)；これで $(k+1,k+1)$ 要素は $a_{i,k+1}^{(k)}$ $(i=k+1,\cdots,m)$ の最大公約数(これを c とする)に等しくなるが，$a_{i,k+1}^{(k)}(i=1,\cdots,k)$ の中に c より小さくないものあるいは負のものがあれば，さらに行 $k+1$ を何倍かしてそのような行 i から引く(あるいは足す)ことにより，$(i,k+1)$ 要素 $(i=1,\cdots,k)$ をすべて c より小さい非負の数にする；このようにして得られる行列を $A^{(k+1)}$ とし，$k:=k+1$ とおいて[2]に戻る．

この算法により得られる行列が $U=A^{(r)}$ である．rank $A=$rank $U=r$ は明ら

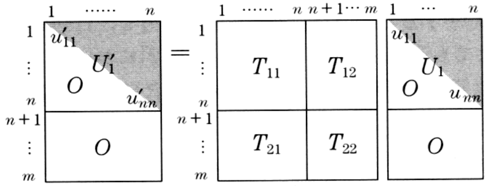
か．

注意 2.15.22 特に rank $A=n$（したがって $m\geqq n$）の場合，rank $A=$rank $A^{(k)}=n$ であるから，$k<n$ のとき，$A^{(k)}$ の $i=k+1,\cdots,m$; $j=k+1,\cdots,n$ の部分の部分行列はどの列にも非零要素がなければならない．したがって，上記算法の[2.2]において，列の入れ換えをする必要がない．このときには，(2.15.88)は

$$SA=U \tag{2.15.91}$$

という形になる．この場合，上記の算法によるかどうかにかかわらず，ある単模行列 S' を用いて $S'A=U'$ が(2.15.89)の性質をもつようにできたとすると，$U=U'$ でなければならないことも，次のようにして示される．

$U'=S'A=S'S^{-1}U$，すなわち $T=S'S^{-1}$ とおくと（S が単模だから S^{-1} も単模，したがって T も単模），



(2.15.92)

という形になるが，U'_1 も U_1 も正則上三角行列で，$U'_1=T_{11}U_1$ であるから T_{11} も上三角行列でなければならない．また，$O=T_{21}U_1$ より $T_{21}=O$．そこで，$|\det T|=|\det T_{11}|\cdot|\det T_{22}|=1$ より $|\det T_{11}|=1$（$\det T_{22}$ は整数）．記号の簡単化のため $T_{11}=B=[b_{ij}]$ とおくと，$|\det T_{11}|=|b_{11}\cdots b_{nn}|=1$，したがって，$|b_{11}|=\cdots=|b_{nn}|=1$．一方，$u'_{ii}=b_{ii}u_{ii}, u'_{ii}>0, u_{ii}>0$ より $b_{ii}=1$．ゆえに，$u'_{ii}=u_{ii}(i=1,\cdots,n)$．次に

$$u'_{j-1,j}=b_{j-1,j-1}u_{j-1,j}+b_{j-1,j}u_{jj}$$

であるが，$b_{j-1,j-1}=1$，$u'_{jj}=u_{jj}$ により，$u'_{j-1,j}-u_{j-1,j}=b_{j-1,j}u_{jj}=b_{j-1,j}u'_{jj}$ となる．このことは，$u_{jj}|(u'_{j-1,j}-u_{j-1,j})$ を意味する．一方，$0\leqq u_{j-1,j}<u_{jj}$，$0\leqq u'_{j-1,j}<u'_{jj}$ であるから，$u'_{j-1,j}=u_{j-1,j}$ でなければならない．すなわち $b_{j-1,j}=0$ でなければならない．すると，

$$u'_{j-2,j}=b_{j-2,j-2}u_{j-2,j}+b_{j-2,j-1}u_{j-1,j}+b_{j-2,j}u_{jj}=u_{j-2,j}+b_{j-2,j}u_{jj}$$

となり，上と同様の議論によって，$b_{j-2,j}=0$ でなければならないことになる．さらに同じ議論を進めることによって，$B=T_{11}=I_n$ でなければならないことがわかる．ゆえに $U'=U$ である．

注意 2.15.23 上記の算法によると，単模行列 A は単模基本行演算によって単模上三角行列 U で(2.15.89)を満たすものに変換されるが，そのような U は単位行列しかない．したがって，単模行列は単模基本行列の積の形に表される．

注意 2.15.24 整数要素の $m\times n$ 型行列 A のすべての k 次小行列式 $|A|^{i_1\cdots i_k}_{j_1\cdots j_k}$ の最

大公約数を $d_k(A)$ と記すことにする．$d_0(A), d_1(A), \cdots, d_l(A)$ $(l=\min(m,n))$ を A の**行列式因子**(determinantal divisors)と呼ぶ．($d_0(A)=1$，$k>\operatorname{rank} A$ に対しては $d_k(A)=0$ と約束する．)すると，Laplace 展開(→2.8.1 項)により，

$$d_k(A)|d_{k+1}(A) \qquad (k \leqq \operatorname{rank} A) \tag{2.15.93}$$

となる．単模基本変形(i′)，(ii′)によって $d_k(A)$ が変化しないことは明らかであるが，(iii′)の型の単模基本変形によっても $d_k(A)$ は変化しない．(実際，変形後の行列の k 次小行列式は変形前のそれらの整数倍の和の形に書けるので，変形後の行列式因子は変形前の因子によって割り切られる．そして，単模基本変形の逆もまた単模基本変形である．）このことと，注意 2.15.23 とを考え合わせると，「S, T が(それぞれ適当な型の)単模行列のとき，A と SAT とは等しい行列式因子をもつ」ことがわかる．

注意 2.15.25　算法を形式的に記述することはもう省くが，(2.15.88)，(2.15.89)の形の変形をするための算法と同様のやり方を，行列 A の行に関してだけでなく，列に関しても適用してみる．すなわち，[2]において，行演算によって列 $k+1$ の(行 $k+2$ 以降の)要素を 0 にしたならば，次に列演算によって行 $k+1$ の(列 $k+2$ 以降の)要素を 0 にする．すると，列 $k+1$ にまた非零が増えるから，再び行演算をおこない，次いで列演算をおこなうということを繰り返す．このような繰り返し 1 回ごとに，$(k+1, k+1)$ 要素の大きさが減少するから，有限回の反復の後，行 $k+1$ と列 $k+1$ の要素は，$(k+1, k+1)$ 要素を除いて，すべて 0 になる．

このようにすれば，結局，S, T を単模行列として，A を

$$SAT = \left[\begin{array}{cccc|c} c_1 & & & O & \\ & c_2 & & & \\ & & \ddots & & O \\ O & & & c_r & \\ \hline & O & & & O \end{array}\right] \qquad (c_i>0) \tag{2.15.94}$$

という形に変形することができる．

さらに，行の置換，列の置換によって，$c_1 < c_2 < \cdots < c_r$ の順に並べられる．

もし $c_1 \nmid c_k$ なる $k(=2, \cdots, r)$ があったら，行 k を行 1 に加えてから再び(2.15.94)の形に変形すると，c_1 が減少する．このようにして，有限回の変形の後，$c_1|c_k$ $(k=2, \cdots, r)$ となる．同様のことを c_2 以降，c_3 以降，…に対しておこなうことによって，結局

$$\tilde{A} = SAT = \left[\begin{array}{ccc|c} e_1 & & & \\ & \ddots & & O \\ & & e_r & \\ \hline & O & & O \end{array}\right]; \quad e_1, \cdots, e_r > 0; \quad e_1|e_2|\cdots|e_r \tag{2.15.95}$$

という形への変形ができる．明らかに

$$d_k(\tilde{A})=e_1\cdot\cdots\cdot e_k \qquad (k=1,\cdots,r) \tag{2.15.96}$$

であるが，$d_k(\tilde{A})=d_k(A)$であることから(注意2.15.24)，$e_1,\cdots,e_r$はS,Tの選び方によらないことがわかる．これらの$e_1,\cdots,e_r$をAの**単因子**(elementary divisors)と呼ぶ．(もちろん，$e_k=d_k(A)/d_{k-1}(A)$である．)

注意 2.15.26　以上のことは6.6節の話の基礎ともなっている重要な事実である．

2.16　特殊な行列と行列式

応用上頻繁に現れる特殊な形をした行列のいくつかについて，それらの特徴的な性質を挙げておこう．

2.16.1　Vandermonde 行列

n個の数(あるいは変数)$x_1,\cdots,x_n$から作られるn次の正方行列

$$V=[v_{ij}], \qquad v_{ij}=x_j^{i-1} \qquad (i,j=1,\cdots,n) \tag{2.16.1}$$

は**Vandermonde**(ヴァンデルモンド)**行列**，det Vは**Vandermonde 行列式**と呼ばれている．(2.16.1)を視覚に訴えるように書くと

$$V=\begin{bmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \\ x_1^2 & x_2^2 & \cdots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1} \end{bmatrix}. \tag{2.16.1$'$}$$

$x_1,\cdots,x_n$を相異なる変数とみなすと(注意2.10.5参照)，det Vは$x_1,\cdots,x_n$の$0+1+\cdots+(n-1)=n(n-1)/2$次の多項式である．これを$x_1$の多項式とみると，ある$i(=2,\cdots,n)$に対して$x_1=x_i$とおくと，$V$は等しい列をもつので det $V=0$となるから，det Vは$(x_1-x_2)(x_1-x_3)\cdots(x_1-x_n)=(-1)^{n-1}(x_2-x_1)\cdots(x_n-x_1)$を因子としてもつ．$x_2,\cdots,x_n$についても同様であるから，det Vは，$x_1,\cdots,x_n$の“基本交代式”(“差積”ともいう)，すなわち

$$\pi(x_1,\cdots,x_n)=\prod_{\substack{i,j=1,\cdots,n \\ i>j}}(x_i-x_j) \tag{2.16.2}$$

を因子としてもつ．ところで，$\pi(x_1,\cdots,x_n)$は，既に$x_1,\cdots,x_n$に関して$n(n$

$-1)/2$ 次であるから，$\det V=c\pi(x_1,\cdots,x_n)$（$c$ は定数）でなければならない．$x_1,\cdots,x_n$ の多項式の一つの項 $x_2x_3{}^2\cdots x_n{}^{n-1}$ の係数を比較すると，$\det V$ の方で $+1$，$\pi(x_1,\cdots,x_n)$ の方でも $+1$ であるから，

$$\det V=\prod_{\substack{i,j=1,\cdots,n\\ i>j}}(x_i-x_j) \tag{2.16.3}$$

である．

V の逆行列 $U=V^{-1}$ も求めてみよう．そのために，まず $x_1,\cdots,x_n$ から x_i を除いた $n-1$ 個の変数の j 次の基本対称式を $s_{ij}(i=1,\cdots,n\,;\,j=0,\cdots,n-1)$ と記し（$s_{i0}=1$ と約束する），変数 t の $n-1$ 次の多項式 $\pi_i(t)$ を

$$\pi_i(t)=\prod_{j=1}^{n}(t-x_j)/(t-x_i)=(t-x_1)\cdots(t-x_{i-1})(t-x_{i+1})\cdots(t-x_n) \tag{2.16.4}$$

と定義する．定義から明らかなように，

$$\pi_i(t)=\sum_{k=1}^{n}(-1)^{n-k}s_{i,n-k}t^{k-1}, \tag{2.16.5}$$

$$\pi_i(x_j)=0 \qquad (j\neq i \text{ のとき}) \tag{2.16.6}$$

である．そこで，

$$U=[u_{ij}],\quad u_{ij}=(-1)^{n-j}s_{i,n-j}/\pi_i(x_i) \qquad (i,j=1,\cdots,n) \tag{2.16.7}$$

により n 次正方行列 U を定義すれば，(2.16.5)，(2.16.6)，(2.16.7) は

$$UV=I \tag{2.16.8}$$

ということを表しているにすぎない（(2.16.5) において $t=x_i$ とおいてみよ）．すなわち $U=V^{-1}$ である．

2.16.2 Gram 行列式

（一般に複素数を要素とする）$n\times m$ 型行列 $A(n\geqq m)$ の第 j 列ベクトルを $\boldsymbol{a}_j(j=1,\cdots,m)$ と記すと，m 次正方行列 $G=A^{\mathrm{H}}A$ の (i,j) 要素は

$$g_{ij}=\boldsymbol{a}_i{}^{\mathrm{H}}\boldsymbol{a}_j \tag{2.16.9}$$

である．$\det G$ を，ベクトル $\boldsymbol{a}_1,\cdots,\boldsymbol{a}_m$ の **Gram**（グラム）**行列式**（Gramian）という．Binet-Cauchy 展開（→2.8.3 項，(2.8.14)）により

$$\det G=\sum\|A|^{i_1\cdots i_m}_{1\cdots m}|^2\geqq 0 \tag{2.16.10}$$

（和は $\{i_1,\cdots,i_m\}\subseteq\{1,\cdots,n\}, i_1<\cdots<i_m$ であるようなすべての $i_1,\cdots,i_m$ について取る）となる．

このことから，

$$\operatorname{rank} A=m \Longleftrightarrow \det G>0 \tag{2.16.11}$$

であることがわかる．$n<m$ なら $\det G=0$(→(2.8.15))．

例 2.16.1 $m=2$ のとき，

$$A=\begin{bmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{bmatrix}$$

と置くと，$G=A^{\mathrm{H}}A$ は

$$G=\begin{bmatrix} \sum_{i=1}^{n}|a_i|^2 & \sum_{i=1}^{n}\overline{a}_i b_i \\ \sum_{i=1}^{n}a_i\overline{b_i} & \sum_{i=1}^{n}|b_i|^2 \end{bmatrix} \tag{2.16.12}$$

であるから，

$$\det G=\left(\sum_{i=1}^{n}|a_i|^2\right)\left(\sum_{i=1}^{n}|b_i|^2\right)-\left|\sum_{i=1}^{n}\overline{a}_i b_i\right|^2 \geqq 0 \tag{2.16.13}$$

は，いわゆる“Cauchy-Schwarz(コーシー-シュヴァルツ)の不等式”を与える．(2.16.13)において等号が成立するのは，rank $A=0$ あるいは 1 のときであるから，すべての $i,j(=1,\cdots,n\,;\,i<j)$ に対して，

$$|A|_{12}^{ij}=a_ib_j-a_jb_i=0$$

すなわち

$$\frac{a_1}{b_1}=\frac{a_2}{b_2}=\cdots=\frac{a_n}{b_n} \tag{2.16.14}$$

のときであり，そしてそのときに限る．

2.16.3 Hilbert 行列

$x_1,\cdots,x_n$，$y_1,\cdots,y_n$ から

$$A=[a_{ij}], \qquad a_{ij}=1/(x_i+y_j) \tag{2.16.15}$$

で定められる n 次正方行列 A を考える．すなわち，

$$A=\begin{bmatrix} \dfrac{1}{x_1+y_1} & \dfrac{1}{x_1+y_2} & \cdots & \dfrac{1}{x_1+y_n} \\ \dfrac{1}{x_2+y_1} & \dfrac{1}{x_2+y_2} & \cdots & \dfrac{1}{x_2+y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \dfrac{1}{x_n+y_1} & \dfrac{1}{x_n+y_2} & \cdots & \dfrac{1}{x_n+y_n} \end{bmatrix}. \tag{2.16.15'}$$

いま，$x_1,\cdots,x_n$，$y_1,\cdots,y_n$ の n^2 次の多項式 $\tilde{\omega}$ を

$$\tilde{\omega}=\prod_{i=1}^{n}\prod_{j=1}^{n}(x_i+y_j) \tag{2.16.16}$$

で定義する．また，n 次対角行列

$$D=\mathrm{diag}\left[\prod_{j=1}^{n}(x_1+y_j),\prod_{j=1}^{n}(x_2+y_j),\cdots,\prod_{j=1}^{n}(x_n+y_j)\right] \tag{2.16.17}$$

を定義する．明らかに

$$\det D=\tilde{\omega} \tag{2.16.18}$$

である．DA の要素はすべて x_i,y_j の $(n-1)$ 次の多項式であるから，$\det(DA)$ は x_i,y_j の $n(n-1)$ 次の多項式である．ある異なる i,j に対して，$x_i=x_j$ とおくと A の行 i と行 j が等しくなるから，$\det A=0$，すなわち $\det(DA)=0$．したがって，$\det(DA)$ は $\pi(x_1,\cdots,x_n)$ (→(2.16.2)) で割り切れる．同様にして，$\det(DA)$ は $\pi(y_1,\cdots,y_n)$ で割り切れる．$\pi(x_1,\cdots,x_n)$ は x_i の $n(n-1)/2$ 次の多項式，$\pi(y_1,\cdots,y_n)$ は y_j の $n(n-1)/2$ 次の多項式であるから，ある定数 c が存在して，

$$\begin{aligned}\det(DA)&=\tilde{\omega}\det A\\ &=c\pi(x_1,\cdots,x_n)\pi(y_1,\cdots,y_n)\end{aligned} \tag{2.16.19}$$

と表せるはずである．c を定めるために，

$$x_k=k,\quad y_k=-k+\varepsilon\qquad(k=1,2,\cdots,n\,;\,|\varepsilon|\ll 1) \tag{2.16.20}$$

とおいてみると，

$$\begin{aligned}&\tilde{\omega}\fallingdotseq(-1)^{n(n-1)/2}\varepsilon^n(2^{n-2}3^{n-3}\cdots(n-2)^2(n-1))^2,\\ &\pi(x_1,\cdots,x_n)=2^{n-2}3^{n-3}\cdots(n-2)^2(n-1),\\ &\pi(y_1,\cdots,y_n)\fallingdotseq(-1)^{n(n-1)/2}2^{n-2}3^{n-3}\cdots(n-2)^2(n-1),\\ &\det A\fallingdotseq\varepsilon^{-n}\end{aligned} \tag{2.16.21}$$

となるので，$c=1$，すなわち，

$$\det A = \pi(x_1,\cdots,x_n)\,\pi(y_1,\cdots,y_n)/\tilde{\omega}$$
$$= \prod_{\substack{i,j=1,\cdots,n\\ i>j}}(x_i-x_j)\prod_{\substack{i,j=1,\cdots,n\\ i>j}}(y_i-y_j)\Big/\prod_{i,j=1,\cdots,n}(x_i+y_j) \quad (2.16.22)$$

であることがわかる．

(2.16.15′)において，特に $x_k=k, y_k=k-1(k=1,\cdots,n)$ とおいたもの，すなわち，

$$B=\begin{bmatrix} 1 & 1/2 & 1/3 & \cdots & 1/n \\ 1/2 & 1/3 & 1/4 & \cdots & 1/(n+1) \\ 1/3 & 1/4 & 1/5 & \cdots & 1/(n+2) \\ \vdots & \vdots & \vdots & & \vdots \\ 1/n & 1/(n+1) & 1/(n+2) & \cdots & 1/(2n-1) \end{bmatrix} \quad (2.16.23)$$

は，n 次の **Hilbert**(ヒルベルト)**行列**と呼ばれている．この行列式は，

$$\tilde{\omega}=n!\frac{(n+1)!}{1!}\frac{(n+2)!}{2!}\cdots\frac{(2n-1)!}{(n-1)!},$$
$$\pi(1,2,\cdots,n)=(n-1)!(n-2)!\cdots2!1!,$$
$$\pi(0,1,\cdots,n-1)=\pi(1,2,\cdots,n)$$

であることを考慮すれば，(2.16.22)より，

$$\det B=\frac{((n-1)!(n-2)!\cdots3!2!)^3}{(2n-1)!(2n-2)!\cdots(n+1)!n!} \quad (2.16.24)$$

と表される．

Hilbert 行列 B の部分行列はみな(2.16.15′)の形をしているから，上と同様のやり方で，その逆行列，LU 分解の要素なども“閉じた形”で表すことができる(→(2.7.7)，(2.7.9)，(2.10.1)，(2.15.35)，(2.15.37)，(2.15.39)，(2.15.43)などを見よ)．

注意 2.16.1　(2.16.15)，(2.16.15′)の形の行列 A に対する公式(2.16.22)を導出するには，上で用いたような高踏的な方法によらないでも，2.6.1 項の行列式の基本的な性質(i)と注意 2.6.2(ハ)(において $\lambda=-1$ とおいたもの)，それに 2.8 節の Laplace 展開を繰り返し用いることによっても可能である．すなわち，A の第 1 列を他の各列から引くと，

$$\frac{1}{x_i+y_j}-\frac{1}{x_i+y_1}=\frac{y_1-y_j}{(x_i+y_j)(x_i+y_1)}$$

であるから，

$$\det A=\det\begin{bmatrix} \dfrac{1}{x_1+y_1} & \dfrac{y_1-y_2}{(x_1+y_2)(x_1+y_1)} & \cdots & \dfrac{y_1-y_n}{(x_1+y_n)(x_1+y_1)} \\ \vdots & \vdots & & \vdots \\ \dfrac{1}{x_n+y_1} & \dfrac{y_1-y_2}{(x_n+y_2)(x_n+y_1)} & \cdots & \dfrac{y_1-y_n}{(x_n+y_n)(x_n+y_1)} \end{bmatrix},$$

各列 $j(=2,\cdots,n)$ の要素には共通の因子 y_1-y_j があり，各行 $i(=1,\cdots,n)$ には共通の因子 $1/(x_i+y_1)$ があるから，

$$\det A=\prod_{i=1}^{n}\frac{1}{x_i+y_1}\prod_{j=2}^{n}(y_1-y_j)\det\begin{bmatrix} 1 & \dfrac{1}{x_1+y_2} & \cdots & \dfrac{1}{x_1+y_n} \\ \vdots & \vdots & & \vdots \\ 1 & \dfrac{1}{x_n+y_2} & \cdots & \dfrac{1}{x_n+y_n} \end{bmatrix}.$$

さらに，第1行を他の各行から引くと，

$$\frac{1}{x_i+y_j}-\frac{1}{x_1+y_j}=\frac{x_1-x_i}{(x_i+y_j)(x_1+y_j)}$$

であるから，列1に関する Laplace 展開の後

$$\det A=\prod_{i=1}^{n}\frac{1}{x_i+y_1}\prod_{j=2}^{n}(y_1-y_j)\cdot\det\begin{bmatrix} \dfrac{x_1-x_2}{(x_2+y_2)(x_1+y_2)} & \cdots & \dfrac{x_1-x_2}{(x_2+y_n)(x_1+y_n)} \\ \vdots & & \vdots \\ \dfrac{x_1-x_n}{(x_n+y_2)(x_1+y_2)} & \cdots & \dfrac{x_1-x_n}{(x_n+y_n)(x_1+y_n)} \end{bmatrix}$$

となるが(右辺の det[…] の次数は $n-1$ になっていることに注意)，さらに各行の共通因子 $x_1-x_i(i=2,\cdots,n)$，各列の共通因子 $1/(x_1+y_j)(j=2,\cdots,n)$ をくくり出すと，

$$\det A=\prod_{i=2}^{n}(x_1-x_i)\prod_{j=2}^{n}(y_1-y_j)\prod_{i=1}^{n}\frac{1}{x_i+y_1}\prod_{j=2}^{n}\frac{1}{x_1+y_j}\det\begin{bmatrix} \dfrac{1}{x_2+y_2} & \cdots & \dfrac{1}{x_2+y_n} \\ \vdots & & \vdots \\ \dfrac{1}{x_n+y_2} & \cdots & \dfrac{1}{x_n+y_n} \end{bmatrix}$$

となって，(2.16.15′) と同じ形の $(n-1)$ 次の正方行列の行列式が残る．同様のことを繰り返せば結局

$$\det A=\prod_{\substack{i,j=1,\cdots,n\\ i<j}}(x_i-x_j)\prod_{\substack{i,j=1,\cdots,n\\ i<j}}(y_i-y_j)\prod_{i,j=1,\cdots,n}\frac{1}{x_i+y_j},$$

すなわち(2.16.22)に到達する($\pi(x_1,\cdots)$ と $\pi(y_1,\cdots)$ との両方に同じ符号 $(-1)^{n(n-1)/2}$ がつくがそれらは消し合うので問題ない)．

注意 2.16.2　(2.16.23)の形の Hilbert 行列 B は数値計算の分野ではきわめて

“性質の悪い”行列として悪名の高いものである．たとえば，標準的な Gauss の消去法(→4.2 節)を用いて，通常の計算環境で，n が大きいとき，その行列式や逆行列を計算したり，B を係数とする連立一次方程式を解いたりしようとすると，たちまち丸め誤差のための桁落ちが生じ，意味のある結果が得られないことが知られている．

2.16.4 Cauchy 行列

Hilbert 型の親戚筋の

$$A=\left[\frac{1}{1-x_iy_j}\right] \tag{2.16.25}$$

という形の行列を **Cauchy 行列**と呼ぶ人がいる．これについても，上の 2.16.3 項(および注意 2.16.1)におけるとほとんど同様のやり方で

$$\det A=\pi(x_1,\cdots,x_n)\,\pi(y_1,\cdots,y_n)\prod_{i,j=1,\cdots,n}\frac{1}{1-x_iy_j} \tag{2.16.26}$$

であることを直接示すことができる．($y_j\leftarrow -1/y_j$ という変数の変換をおこなって 2.16.3 項の場合に帰着させることもできる．)

2.16.5 Hadamard 行列

$$H_2=\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \tag{2.16.27}$$

は

$$H_2{}^{\mathsf T}=H_2,\quad H_2\cdot H_2=2I_2 \tag{2.16.28}$$

を満たす．すなわち，$2^{-1/2}H_2$ は対称な直交行列である．ある $p(\geqq 1)$ に対して，H_{2^p} が 2^p 次正方行列で

$$H_{2^p}{}^{\mathsf T}=H_{2^p},\quad H_{2^p}\cdot H_{2^p}=2^pI_{2^p} \tag{2.16.29}$$

を満たせば，2^{p+1} 次正方行列

$$H_{2^{p+1}}=\begin{bmatrix}H_{2^p} & H_{2^p}\\ H_{2^p} & -H_{2^p}\end{bmatrix} \tag{2.16.30}$$

はやはり(2.16.29)の形の関係式(p の代りに $p+1$ とおいた式)を満たす．

一般に，要素が 1 および -1 の n 次の対称な正方行列 H_n が

$$H_n\cdot H_n=nI_n \tag{2.16.31}$$

を満たすとき，H_n は **Hadamard**(アダマール)**行列**と呼ばれる．$n=2^p$ ($p=1,2,\cdots$) に対して Hadamard 行列が存在することは，上のようにして，容易に知られる．それ以外の n に対しては，Hadamard 行列の次数は 4 の倍数でなければならないことが知られているが，すべての 4 の倍数に対してその次数の Hadamard 行列が存在するかどうかはわかっていない．(Hadamard 行列は，実験計画法，符号理論などに関連した“組合せ論”の問題として研究されている．)

2.16.6　Hankel 行列と Toeplitz 行列

数列 $u=(\cdots,u^{(-2)},u^{(-1)},u^{(0)},u^{(1)},u^{(2)},\cdots)$ (両方向に無限に続くものととりあえずは考えておく) に関連した“$u^{(\nu)}$を基準とする k 次の” **Hankel**(ハンケル)**行列**は，

$$H_k^{(\nu)}(u)=\begin{bmatrix} u^{(\nu)} & u^{(\nu-1)} & \cdots & u^{(\nu-k+1)} \\ u^{(\nu-1)} & u^{(\nu-2)} & \cdots & u^{(\nu-k)} \\ \vdots & \vdots & \iddots & \vdots \\ u^{(\nu-k+1)} & u^{(\nu-k)} & \cdots & u^{(\nu-2k+2)} \end{bmatrix} \tag{2.16.32}$$

と定義される．(右上から左下方向へ斜めに等しい要素が並ぶ) この形の行列は，繰り返し構造をもつシステムや計算過程の解析，それに時系列解析などに関連してよく現れる．

数列 u の**後退差分**(backward difference) $v=\nabla u$ とは，

$$v^{(\nu)}=u^{(\nu)}-u^{(\nu-1)} \qquad (\nu=\cdots,-1,0,1,\cdots) \tag{2.16.33}$$

で定義される数列 $v=(\cdots,v^{(-1)},v^{(0)},v^{(1)},\cdots)$ のことである．$\nabla^2 u=\nabla(\nabla u)$ なども同様に定義される．

いま，$k+1$ 次の行列式

$$a=\det\begin{bmatrix} u^{(\nu)} & \cdots & u^{(\nu-k+1)} & u^{(\nu-k)} \\ \vdots & \iddots & \vdots & \vdots \\ u^{(\nu-k+1)} & \cdots & u^{(\nu-2k+2)} & u^{(\nu-2k+1)} \\ 1 & \cdots & 1 & 1 \end{bmatrix} \tag{2.16.34}$$

の値を計算するのに，第 1 列から第 2 列を引き，第 2 列から第 3 列を引き，…，第 k 列から第 $k+1$ 列を引いてみると，

$$a=\det\begin{bmatrix} \nabla u^{(\nu)} & \cdots & \nabla u^{(\nu-k+1)} & u^{(\nu-k)} \\ \vdots & \iddots & \vdots & \vdots \\ \nabla u^{(\nu-k+1)} & \cdots & \nabla u^{(\nu-2k+2)} & u^{(\nu-2k+1)} \\ 0 & \cdots & 0 & 1 \end{bmatrix} \tag{2.16.35}$$

となることから，

$$a=\det H_k^{(\nu)}(\nabla u) \tag{2.16.36}$$

であることがわかる．同様にして，$k+2$ 次の行列式

$$b=\det\begin{bmatrix} u^{(\nu)} & \cdots & u^{(\nu-k+1)} & u^{(\nu-k)} & 1 \\ \vdots & \iddots & \vdots & \vdots & \vdots \\ u^{(\nu-k+1)} & \cdots & u^{(\nu-2k+2)} & u^{(\nu-2k+1)} & 1 \\ u^{(\nu-k)} & \cdots & u^{(\nu-2k+1)} & u^{(\nu-2k)} & 1 \\ 1 & \cdots & 1 & 1 & 0 \end{bmatrix} \tag{2.16.37}$$

は，

$$b=\det\begin{bmatrix} \nabla^2 u^{(\nu)} & \cdots & \nabla^2 u^{(\nu-k+1)} & \nabla u^{(\nu-k)} & 0 \\ \vdots & \iddots & \vdots & \vdots & \vdots \\ \nabla^2 u^{(\nu-k+1)} & \cdots & \nabla^2 u^{(\nu-2k+2)} & \nabla u^{(\nu-2k+1)} & 0 \\ \nabla u^{(\nu-k)} & \cdots & \nabla u^{(\nu-2k+1)} & u^{(\nu-2k)} & 1 \\ 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \tag{2.16.38}$$

と変形されるから，

$$b=-\det H_k^{(\nu)}(\nabla^2 u) \tag{2.16.39}$$

となる．

$\det H_k^{(\nu)}(u)$, $\det H_k^{(\nu)}(\nabla u)$, $\det H_k^{(\nu)}(\nabla^2 u)$ は，ε-算法，商差法と呼ばれる算法などにおいて重要な役割を果たす多くの有用な漸化式を満たす．たとえば，

$$\det\begin{bmatrix} u^{(\nu)} & u^{(\nu-1)} & \cdots & u^{(\nu-k+1)} & u^{(\nu-k)} & 1 \\ u^{(\nu-1)} & u^{(\nu-2)} & \cdots & u^{(\nu-k)} & u^{(\nu-k-1)} & 1 \\ \vdots & \vdots & \iddots & \vdots & \vdots & \vdots \\ u^{(\nu-k+1)} & u^{(\nu-k)} & \cdots & u^{(\nu-2k+2)} & u^{(\nu-2k+1)} & 1 \\ u^{(\nu-k)} & u^{(\nu-k-1)} & \cdots & u^{(\nu-2k+1)} & u^{(\nu-2k)} & 1 \\ 1 & 1 & \cdots & 1 & 1 & 0 \end{bmatrix} \tag{2.16.40}$$

に対して，2.10.4 項の Sylvester の定理(xvii)(式(2.10.28)，(2.10.29))を適用

すると($\boxed{}$の部分を(xvii)の A とみなす；また上の(2.16.36)，(2.16.39)も利用する)，

(i)　$$(\det H_k^{(\nu)}(\nabla^2 u))(\det H_k^{(\nu-1)}(u)) = (\det H_{k-1}^{(\nu-1)}(\nabla^2 u))(\det H_{k+1}^{(\nu)}(u)) + (\det H_k^{(\nu)}(\nabla u))(\det H_k^{(\nu-1)}(\nabla u)) \tag{2.16.41}$$

を得る．また，

$$\det\begin{bmatrix} u^{(\nu)} & u^{(\nu-1)} & \cdots & u^{(\nu-k)} & 1 & 1 \\ u^{(\nu-1)} & u^{(\nu-2)} & \cdots & u^{(\nu-k-1)} & 1 & 0 \\ \vdots & \vdots & \iddots & \vdots & \vdots & \vdots \\ u^{(\nu-k)} & u^{(\nu-k-1)} & \cdots & u^{(\nu-2k)} & 1 & 0 \\ u^{(\nu-k-1)} & u^{(\nu-k-2)} & \cdots & u^{(\nu-2k-1)} & 1 & 0 \\ 1 & 1 & \cdots & 1 & 0 & 0 \end{bmatrix} \tag{2.16.42}$$

を考えると，

(ii)　$$(\det H_{k+1}^{(\nu)}(u))(\det H_k^{(\nu-1)}(\nabla^2 u)) = (\det H_{k+1}^{(\nu)}(\nabla u))(\det H_k^{(\nu-1)}(\nabla u)) + (\det H_k^{(\nu)}(\nabla^2 u))(\det H_{k+1}^{(\nu-1)}(u)) \tag{2.16.43}$$

を得る．さらに，

$$\det\begin{bmatrix} u^{(\nu)} & u^{(\nu-1)} & \cdots & u^{(\nu-k+1)} & u^{(\nu-k)} & 1 \\ u^{(\nu-1)} & u^{(\nu-2)} & \cdots & u^{(\nu-k)} & u^{(\nu-k-1)} & 1 \\ \vdots & \vdots & \iddots & \vdots & \vdots & \vdots \\ u^{(\nu-k+1)} & u^{(\nu-k)} & \cdots & u^{(\nu-2k+2)} & u^{(\nu-2k+1)} & 1 \\ u^{(\nu-k)} & u^{(\nu-k-1)} & \cdots & u^{(\nu-2k+1)} & u^{(\nu-2k)} & 1 \\ 1 & 1 & \cdots & 1 & 1 & 0 \end{bmatrix} \tag{2.16.44}$$

の2通りを考えると，

(iii)　$$(\det H_k^{(\nu)}(u))(\det H_k^{(\nu)}(\nabla^2 u)) = (\det H_{k+1}^{(\nu)}(u))(\det H_{k-1}^{(\nu)}(\nabla^2 u)) + (\det H_k^{(\nu)}(\nabla u))^2, \tag{2.16.45}$$

(iv)　$$(\det H_k^{(\nu-2)}(u))(\det H_k^{(\nu)}(\nabla^2 u)) = (\det H_{k+1}^{(\nu)}(u))(\det H_{k-1}^{(\nu-2)}(\nabla^2 u)) + (\det H_k^{(\nu-1)}(\nabla u))^2 \tag{2.16.46}$$

が得られる．

数列 u の差分を含まないものとしては，

$$\det\begin{bmatrix} u^{(\nu)} & u^{(\nu-1)} & \cdots & u^{(\nu-k+2)} & u^{(\nu-k+1)} \\ u^{(\nu-1)} & u^{(\nu-2)} & \cdots & u^{(\nu-k+1)} & u^{(\nu-k)} \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ u^{(\nu-k+2)} & u^{(\nu-k+1)} & \cdots & u^{(\nu-2k+4)} & u^{(\nu-2k+3)} \\ u^{(\nu-k+1)} & u^{(\nu-k)} & \cdots & u^{(\nu-2k+3)} & u^{(\nu-2k+2)} \end{bmatrix} \tag{2.16.47}$$

を考えることにより，

(v)　$(\det H_{k-2}^{(\nu-2)}(u))(\det H_k^{(\nu)}(u))$

$$=(\det H_{k-1}^{(\nu)}(u))(\det H_{k-1}^{(\nu-2)}(u))-(\det H_{k-1}^{(\nu-1)}(u))^2 \tag{2.16.48}$$

が得られる．

信号処理論，確率過程論などに関連して重要な **Toeplitz**(テプリッツ)**行列**は，数列 $v=(v^{(0)},v^{(1)},v^{(2)},\cdots)$ に関連した

$$T_n=T_n(v)=\begin{bmatrix} v^{(0)} & v^{(1)} & \cdots & v^{(n-2)} & v^{(n-1)} \\ v^{(1)} & v^{(0)} & \ddots & & v^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ v^{(n-2)} & & \ddots & v^{(0)} & v^{(1)} \\ v^{(n-1)} & v^{(n-2)} & \cdots & v^{(1)} & v^{(0)} \end{bmatrix} \tag{2.16.49}$$

の形の n 次の対称行列，すなわち，

$$T_n=[t_{ij}^{(n)}],\quad t_{ij}^{(n)}=v^{(|i-j|)} \tag{2.16.49'}$$

で定義される．Toeplitz 行列は，行(あるいは列)の並び順を逆転すれば Hankel 行列の特殊な場合に帰着するので，Toeplitz 行列も Hankel 行列と同様多くの有用な漸化関係式を満たす．

やや技巧的になるが，T_n の第 k 列を縦ベクトル

$$[v^{(1)},v^{(2)},\cdots,v^{(n)}]^{\mathrm{T}} \tag{2.16.50}$$

でおきかえた行列を

$$T_n^k=\begin{bmatrix} v^{(0)} & \cdots & v^{(k-2)} & v^{(1)} & v^{(k)} & \cdots & v^{(n-1)} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ v^{(n-1)} & \cdots & v^{(n-k+1)} & v^{(n)} & v^{(n-k-1)} & \cdots & v^{(0)} \end{bmatrix} \tag{2.16.51}$$

と記す．

Hankel 行列のときと同様に，

$$\det\begin{bmatrix} v^{(0)} & v^{(1)} & \cdots & \cdots & v^{(n-1)} & v^{(n)} \\ v^{(1)} & v^{(0)} & v^{(1)} & \cdots & v^{(n-2)} & v^{(n-1)} \\ v^{(2)} & v^{(1)} & v^{(0)} & \ddots & \vdots & v^{(n-2)} \\ \vdots & \vdots & \ddots & \ddots & v^{(1)} & \vdots \\ v^{(n-1)} & v^{(n-2)} & \cdots & v^{(1)} & v^{(0)} & v^{(1)} \\ v^{(n)} & v^{(n-1)} & \cdots & \cdots & v^{(1)} & v^{(0)} \end{bmatrix} \tag{2.16.52}$$

を考えて，Sylvester の公式を用いると，

$$\text{(vi)}\quad (\det T_{n-1})(\det T_{n+1}) = (\det T_n)^2 - (\det T_n^n)^2 \tag{2.16.53}$$

が得られる．$(\det T_n)\cdot(\det T_{n-1}) \neq 0$ ならば，この式は

$$\text{(vi}'\text{)}\quad \frac{\det T_{n+1}}{\det T_n} \Big/ \frac{\det T_n}{\det T_{n-1}} = 1 - \left(\frac{\det T_n^n}{\det T_n}\right)^2 \tag{2.16.53$'$}$$

とも書ける．また，

$$\det\begin{bmatrix} \overset{0}{0} & \overset{1}{0} & \cdots & 0 & \overset{k}{1} & 0 & \cdots & \overset{n-1}{0} & \overset{n}{0} \\ v^{(1)} & v^{(0)} & v^{(1)} & \cdots & \cdots & \cdots & \cdots & v^{(n-2)} & v^{(n-1)} \\ v^{(2)} & v^{(1)} & \ddots & & & & & \vdots & v^{(n-2)} \\ \vdots & \vdots & & \ddots & & & & \vdots & \vdots \\ \vdots & \vdots & & & \ddots & & & \vdots & \vdots \\ \vdots & \vdots & & & & \ddots & & v^{(1)} & \vdots \\ v^{(n-1)} & v^{(n-2)} & \cdots & \cdots & \cdots & \cdots & v^{(1)} & v^{(0)} & v^{(1)} \\ v^{(n)} & v^{(n-1)} & v^{(n-2)} & \cdots & \cdots & \cdots & \cdots & v^{(1)} & v^{(0)} \end{bmatrix} \tag{2.16.54}$$

を考えると，

$$\text{(vii)}\quad (\det T_{n-1})(\det T_n^k) = (\det T_n)(\det T_{n-1}^k) - (\det T_n^n)(\det T_{n-1}^{n-k}), \tag{2.16.55}$$

あるいは，

$$\text{(vii}'\text{)}\quad \frac{\det T_n^k}{\det T_n} = \frac{\det T_{n-1}^k}{\det T_{n-1}} - \frac{\det T_n^n}{\det T_n}\cdot\frac{\det T_{n-1}^{n-k}}{\det T_{n-1}} \tag{2.16.55$'$}$$

が得られる．(なお，(2.16.55)の右辺の最後尾の行列式は，もともと

$$\det\left[\begin{array}{ccc:c:cc} v^{(0)} & \cdots & v^{(k-2)} & v^{(n-1)} & v^{(k)} & \cdots v^{(n-2)} \\ \vdots & & \vdots & \vdots & \vdots & \vdots \\ v^{(n-2)} & \cdots & v^{(n-k)} & v^{(1)} & v^{(n-k-2)} \cdots & v^{(0)} \end{array}\right] \tag{2.16.56}$$

の形をしたものであるが，行の並びを逆転し，そして列の並びも逆転してみればわかるように，これは $\det T_{n-1}^{n-k}$ に等しい．)

2.16.7　巡回行列

$$C=\begin{bmatrix} c_0 & c_1 & c_2 & \cdots & c_{n-2} & c_{n-1} \\ c_{n-1} & c_0 & c_1 & \cdots & c_{n-3} & c_{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ c_2 & c_3 & c_4 & \cdots & c_0 & c_1 \\ c_1 & c_2 & c_3 & \cdots & c_{n-1} & c_0 \end{bmatrix} \tag{2.16.57}$$

すなわち，

$$C=[c_{ij}]:\qquad c_{ij}=c_k\quad (k=j-i \bmod n) \tag{2.16.58}$$

の形の行列を**巡回行列**(circulant(matrix))という．

例 2.14.7 のユニタリ行列 Z (→(2.14.36))を用いて，ベクトル

$$\boldsymbol{c}=[c_0, c_1, \cdots, c_{n-1}]^{\mathsf T}$$

の離散 Fourier 変換を

$$\boldsymbol{\gamma}=[\gamma_0, \gamma_1, \cdots, \gamma_{n-1}]^{\mathsf T}=Z\boldsymbol{c},\qquad \gamma_{i-1}=\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} c_j\zeta^{j(i-1)} \tag{2.16.59}$$

とすると，$\boldsymbol{\gamma}^{\mathsf T}=[\gamma_0, \gamma_1, \cdots, \gamma_{n-1}]=\boldsymbol{c}^{\mathsf T}Z^{\mathsf T}=\boldsymbol{c}^{\mathsf T}Z$ であるから，CZ の(k, l)要素は

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} c_{j-k+1}\zeta^{j(l-1)} &= \frac{1}{\sqrt{n}}\sum_{j=0}^{n-1} c_j\zeta^{(j+k-1)(l-1)} \\ &= \frac{1}{\sqrt{n}}\zeta^{(k-1)(l-1)}\sum_{j=0}^{n-1} c_j\zeta^{j(l-1)} \\ &= \zeta^{(k-1)(l-1)}\gamma_{l-1} \end{aligned} \tag{2.16.60}$$

となる．すなわち，

$$CZ=\sqrt{n}Z\,\mathrm{diag}[\gamma_0, \gamma_1, \cdots, \gamma_{n-1}] \tag{2.16.61}$$

となる．したがって，

$$Z^{\mathsf H}CZ=\sqrt{n}Z^{\mathsf H}Z\,\mathrm{diag}[\gamma_0, \gamma_1, \cdots, \gamma_{n-1}]=\sqrt{n}\,\mathrm{diag}[\gamma_0, \gamma_1, \cdots, \gamma_{n-1}] \tag{2.16.62}$$

である．

式(2.16.62)は巡回行列の特徴をよく表している．

たとえば，$\det(Z^{\mathrm{H}}CZ)=(\det Z^{\mathrm{H}})(\det C)(\det Z)=\det C$ であるから，

$$\det C = n^{n/2}\gamma_0\cdot\gamma_1\cdot\cdots\cdot\gamma_{n-1}=\prod_{k=0}^{n-1}(\sqrt{n}\gamma_k) \tag{2.16.63}$$

である．

また，C と D が共に n 次の巡回行列であれば $CD=DC$ である．実際，$Z^{\mathrm{H}}CZ$ も $Z^{\mathrm{H}}DZ$ も対角行列であるから $(Z^{\mathrm{H}}CZ)(Z^{\mathrm{H}}DZ)=(Z^{\mathrm{H}}DZ)(Z^{\mathrm{H}}CZ)$，すなわち $Z^{\mathrm{H}}CDZ=Z^{\mathrm{H}}DCZ$，ゆえに $CD=DC$ となる．このことは，同じ型の巡回行列は行列積に関して可換であることを示している．

例 2.16.2　$\zeta=(-1+\sqrt{3}\,\mathrm{i})/2$ は，1 の原始 3 乗根である．具体的には，$\zeta^2=(-1-\sqrt{3}\,\mathrm{i})/2$，$\zeta^3=1$，$\zeta+\zeta^2=-1$．$[a,b,c]^{\mathrm{T}}$ の離散 Fourier 変換は

$$Z\begin{bmatrix}a\\b\\c\end{bmatrix}=\frac{1}{\sqrt{3}}\begin{bmatrix}a+b+c\\a+\zeta b+\zeta^2 c\\a+\zeta^2 b+\zeta c\end{bmatrix} \tag{2.16.64}$$

であるから，(2.16.63)により，

$$\begin{aligned}\det\begin{bmatrix}a&b&c\\c&a&b\\b&c&a\end{bmatrix}&=a^3+b^3+c^3-3abc\\&=(a+b+c)(a+\zeta b+\zeta^2 c)(a+\zeta^2 b+\zeta c)\\&=(a+b+c)(a^2+b^2+c^2-bc-ca-ab)\end{aligned} \tag{2.16.65}$$

を得る．　□

2.17　パーマネント

行列式と似て非なる概念であるが，行列式と多くの奇妙な関係をもち，かつ組合せ数学などによく現れる“パーマネント”についてごく簡単に触れておこう．

2.17.1　パーマネントの定義と基本的性質

n 次正方行列 $A=[a_{ij}]$ の**パーマネント**(permanent)は

$$\mathrm{per}\,A = \sum a_{1j_1}a_{2j_2}\cdots a_{nj_n} \tag{2.17.1}$$

($\sum$ は$(1\cdots n)$の $n!$ 個のすべての置換$(j_1\cdots j_n)$にわたって取る)で定義される．行列式の定義(2.6.1)と比較すれば，$\sum$ の中の各項に符号 $\varepsilon^{1\,2\,\cdots n}_{j_1j_2\cdots j_n}$ が付いていない点だけが異なる．

注意 2.17.1 行列式に関連して，交代化演算(→2.5節)が重要な役割を果たしたが，r 個の添字 $i_1\cdots i_r$ がこの順に並んでいる式 $F_{i_1\cdots i_r}$ に対して，これらの添字に関して $F_{i_1\cdots i_r}$ を**対称化**する(to symmetrize)ということを

$$F_{(i_1\cdots i_r)} = \frac{1}{r!}\sum F_{j_1\cdots j_r} \tag{2.17.2}$$

を作ることと定義する($\sum$ は$(i_1\cdots i_r)$の $r!$ 個のすべての置換$(j_1\cdots j_r)$にわたって取る)．$F_{i_1\cdots i_r} = F_{(i_1\cdots i_r)}$ のとき，$F_{i_1\cdots i_r}$ は $i_1,\cdots,i_r$ に関して**対称**(symmetric)であるという．

対称化演算は交代化演算と同様の性質をもつ．2.5.2項で述べた交代化演算の諸性質は，符号を除いて(すなわち“$-$”は“$+$”でおきかえて)すべて成り立つ．ただ一つ，重大な差異がある．それは，$i_1,\cdots,i_r$ の中に同じ値のものが2個以上あるときには，$F_{[i_1\cdots i_r]}=0$ であるが，$F_{(i_1\cdots i_r)}$ は0になるとは限らないということである．

対称化演算の形式を用いると，行列式の表現(2.6.2)にならって，パーマネントも

$$\mathrm{per}\,A = n!\ a^{1}_{(1}a^{2}_{2}\cdots a^{n}_{n)} = n!\ a^{(1}_{1}a^{2}_{2}\cdots a^{n)}_{n} = n!\ a^{(1}_{(1}a^{2}_{2}\cdots a^{n)}_{n)} \tag{2.17.3}$$

と表せる．しかし，

$$\det A = \sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n} a^{[i_1}_{[i_1}\cdots a^{i_n]}_{i_n]} \tag{2.17.4}$$

ではあるが，一般には，

$$\mathrm{per}\,A = \sum_{i_1=1}^{n}\cdots\sum_{i_n=1}^{n} a^{(i_1}_{(i_1}\cdots a^{i_n)}_{i_n)} \tag{2.17.5}$$

とはならない．

パーマネントは行列式と類似の性質を多くもつ．2.6.1項の行列式の性質と比較して述べると，注意2.6.1(式(2.6.3))，(i)～(iv)(式(2.6.4)～(2.6.7))はそのまま((iii)(2.6.6)については“$-$”を“$+$”にして)成り立つ．しかし，注意2.6.2の(イ)(2.6.8)は成り立つが，(ロ)(2.6.9)，(ハ)(2.6.10)は成り立た**ない．**

一般に $m\times n$ 型行列 $A=[a_{ij}]$ の要素から作られる r 次の正方行列 $B=[b_{\alpha\beta}]$

$(r \leqq m, n)$ で $b_{\alpha\beta}=a_{i_\alpha j_\beta}(\alpha, \beta=1,\cdots,r)$ であるものを，$B=(A)^{i_1\cdots i_r}_{j_1\cdots j_r}$ と書くことにする．ただし，$i_1 \leqq \cdots \leqq i_r$，$j_1 \leqq \cdots \leqq j_r$ と約束する．(ここで，$i_1,\cdots,i_r$ の中にも，$j_1,\cdots,j_r$ の中にも，同じ値のものがあってもよいことにする．この点，2.4.1 項の部分行列の定義とは異なることに注意！)すると，n 次正方行列 A の行列式の Laplace 展開(2.8.6)に相当するものとして，たとえば，

$$\operatorname{per} A=\sum \operatorname{per}(A)^{1\cdots r}_{j_1\cdots j_r} \operatorname{per}(A)^{r+1,\cdots,n}_{j_{r+1}\cdots j_n} \tag{2.17.6}$$

($\sum$ は $\{1,\cdots,n\}$ を r 元部分集合 $\{j_1,\cdots,j_r\}$ とその補集合 $\{j_{r+1},\cdots,j_n\}$ に分割する $\binom{n}{r}=\binom{n}{n-r}$ 通りの分割のすべてにわたって取る)が成り立つ．また，$m \times n$ 型行列 A と $n \times m$ 型行列 B ($m \leqq n$ に限らなくてよい)との積 AB の行列式の Binet-Cauchy 展開(→ 2.8.3 項)に相当するものとして，

$$\operatorname{per}(AB)=\sum \frac{1}{\mu(j_1,\cdots,j_m)} \operatorname{per}(A)^{1\cdots m}_{j_1\cdots j_m} \operatorname{per}(B)^{j_1\cdots j_m}_{1\cdots m} \tag{2.17.7}$$

が成り立つ．ここで，$\sum$ は $\{1,\cdots,n\}$ から**重複を許して** m 個の元 $\{j_1,\cdots,j_m\}$ を選ぶ $\binom{n+m-1}{m}$ 通りの選び方すべてにわたって取る．また，$\mu(j_1,\cdots,j_m)$ は $\{j_1,\cdots,j_m\}$ の中に現れる異なる数それぞれの出現回数の階乗の積である(たとえば，$\mu(2,4,4,4,5,5,7)=1!3!2!1!=12$)．

要素がパラメタ t の関数である行列 $A(t)$ のパーマネントの t に関する導関数 $\dfrac{\mathrm{d}}{\mathrm{d}t}\operatorname{per} A(t)$ に対しても，対応する行列式の場合の式(2.12.18)と同じ形の式が成り立つ．(Laplace 展開に相当する式から導かれる．)

以上の諸性質は，対応する行列式の性質を行列式の定義(2.6.1)から直接導き出したのとほとんど同じ考え方により，導き出すことができる．

注意 2.17.2　明らかに，2 次正方行列 $A=\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$ に対しては

$$\operatorname{per}\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}=\det\begin{bmatrix} a_{11} & -a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

であるが，$n \geqq 3$ に対して，n 次正方行列 A の要素の符号を適当に変えた行列 $T(A)$ (あるいはもっと一般に A の要素の 1 次式を要素とする適当な行列 $T(A)$)を

$$\operatorname{per} A=\det T(A) \tag{2.17.8}$$

となるように作ることは不可能である，しかし，A をもとにもっと大きな行列 B を作って $\operatorname{per} A=\det B$ とすることは可能であることが知られている．

注意 2.17.3 n 次正方行列 A のパーマネント $\mathrm{per}\,A$ の計算には，A の LU 分解を用いる行列式 $\det A$ の計算のような，有効な手順は知られていない．定義式 (2.17.1) や Laplace 展開のような方法によらざるを得ない．また，以下のような計算法も有用な場合がある．

n 次正方行列 $B=[b_{ij}]$ に対して，その各行の要素の和の積を

$$S(B)=\prod_{i=1}^{n}\sum_{j=1}^{n}b_{ij} \tag{2.17.9}$$

と記すことにする．また，A の列 $k_1,k_2,\cdots,k_r\,(k_1<k_2<\cdots<k_r)$ を零ベクトルでおきかえた行列を $A\{k_1,k_2,\cdots,k_r\}$ と記すことにする．すると，

$$S(A)=\sum a_{1j_1}\cdots a_{nj_n} \tag{2.17.10}$$

であって，$\sum$ は $j_1,\cdots,j_n=1,\cdots,n$ のすべてにわたる．$\sum$ の中の項のうち，$j_1,\cdots,j_n$ が互いに異なる場合の和が $\mathrm{per}\,A$ である．(2.17.10) の右辺の $\sum$ の中の項で $\mathrm{per}\,A$ に含まれないものは，$j_1,\cdots,j_n$ の中に等しいものがある，すなわち，$j_1,\cdots,j_n$ の中に $1,\cdots,n$ のすべては現れない．$j_1,\cdots,j_n$ の中に 1 が現れないような項の和は $S(A\{1\})$ と表せる．同様にして，$j_1,\cdots,j_n$ の中に $k_1,\cdots,k_r\,(k_1<\cdots<k_r)$ が現れないような項の和は $S(A\{k_1,\cdots,k_r\})$ と表される．そこで，組合せ数学における基本的な数え上げの手法である "包除原理(principle of inclusion and exclusion)"（たとえば，伊理・藤重・大山：「グラフ・ネットワーク・マトロイド」，産業図書，1986，2005 など参照）によれば，

$$\begin{aligned}\mathrm{per}\,A=S(A)-\textstyle\sum_1 S(A\{k_1\})+\sum_2 S(A\{k_1,k_2\})-\cdots\\ +(-1)^{n-1}\textstyle\sum_{n-1}S(A\{k_1,\cdots,k_{n-1}\})\end{aligned} \tag{2.17.11}$$

が得られる．ここで $\sum_r$ は $\{1,\cdots,n\}$ の中から異なる r 個の元 $\{k_1,\cdots,k_r\}$ を選ぶ $\binom{n}{r}$ 通りの選び方すべてにわたって取る．

例 2.17.1 $\mathrm{per}\,I_n=1$ であって，パーマネントは行列の行の置換，列の置換に関して不変であるから，置換行列 P に対しては $\mathrm{per}\,P=1$ である．

すべての要素が 1 に等しい n 次正方行列を J_n と記すと

$$\mathrm{per}\,J_n=n! \tag{2.17.12}$$

も明らかであろう．

さらに，$\mathrm{per}(J_n+tI_n)$ を t の多項式とみなすとき，次の式もたやすく導き出せる．

$$\mathrm{per}(J_n+tI_n)=\sum_{r=0}^{n}\binom{n}{r}(n-r)!\,t^r=n!\sum_{r=0}^{n}\frac{t^r}{r!} \tag{2.17.13}$$

ここで，$t=-1$ とおくと，

$$\mathrm{per}(J_n - I_n) = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - + (-1)^n \frac{1}{n!}\right) \tag{2.17.14}$$

である．$\mathrm{per}(J_n - I_n)$は，その定義により，「n個のもの$1,\cdots,n$に番号$j_1,\cdots,j_n$ $(=1,\cdots,n)$を付けるとき，i番目のものに番号iが付けられないような(すなわち，$j_i \neq i\,(i=1,\cdots,n)$)番号の付け方が何通りあるか」という問題(Montmort (モンモール)の“**遭遇の問題**(problème des rencontres)”とか“**乱列**(derangements)”の問題とか呼ばれる)の答を与えるものである．

0および1のみを要素とするn次正方行列Aに対し，頂点集合$U=\{u_1,\cdots,u_n\}$，$V=\{v_1,\cdots,v_n\}$をもつ2部グラフで，$a_{ij}=1$のとき，そしてそのときに限りu_iとv_jを結ぶ辺を有するようなものGを作ると，per A＝“Gの上の完全マッチングの総数”であることも明らかであろう(パーマネントの定義の言い換えにすぎない)．a_{ij}が0,1に限らない場合にも，同様な2部グラフを考えて，頂点u_iとv_jを結ぶ辺に“重み”a_{ij}を付与すれば，per A＝“Gの上の完全マッチングの辺の重みの積の和”ということができる． □

例 2.17.2 $P_n=[p_{ij}]$を，$p_{i,i+1}=1\,(i=1,\cdots,n-1)$，$p_{n,1}=1$，その他の$p_{ij}=0$なる置換行列とする：

$$P_n = \begin{bmatrix} 0 & 1 & & & O \\ 0 & 0 & 1 & & \\ \vdots & & & \ddots & \\ 0 & & O & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}. \tag{2.17.15}$$

もちろん，$P_n{}^n = I_n$, per $P_n{}^r = 1\,(r=1,2,\cdots)$であり，さらに，

$$\mathrm{per}(P_n + P_n{}^2) = \mathrm{per}\begin{bmatrix} 0 & 1 & 1 & & O \\ \vdots & \ddots & \ddots & \ddots & \\ 0 & & & \ddots & 1 \\ 1 & 0 & & \ddots & 1 \\ 1 & 1 & 0 & \cdots & 0 \end{bmatrix} = 2 \tag{2.17.16}$$

である．また，

$$f_n = \mathrm{per}(P_n + P_n{}^2 + P_n{}^3) \tag{2.17.17}$$

とおくと，

$$f_n=f_{n-1}+f_{n-2}-2=\left(\frac{1+\sqrt{5}}{2}\right)^n+\left(\frac{1-\sqrt{5}}{2}\right)^n+2=\mathrm{tr}(C_2{}^n)+2,$$

$$C_2=\begin{bmatrix}0 & 1\\ 1 & 1\end{bmatrix} \tag{2.17.18}$$

であること，そして，

$$g_n=\mathrm{per}(P_n+P_n{}^2+P_n{}^3+P_n{}^4) \tag{2.17.19}$$

とおくと，

$$g_n=g_{n-1}+g_{n-2}+g_{n-3}-4=2(\mathrm{tr}(C_3{}^n)+1),$$

$$C_3=\begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 1 & 1\end{bmatrix} \tag{2.17.20}$$

であることが知られている．$r>4$ のときには，$\mathrm{per}\left(\sum_{k=1}^{r}P_n{}^k\right)$に対する同種の漸化式が**存在しない**ことも知られている． □

2.17.2 van der Waerden の予想

n 次正方行列 $A=[a_{ij}]$は

$$a_{ij}\geqq 0 \qquad (i,j=1,\cdots,n),$$

$$\sum_{i=1}^{n}a_{ij}=1 \qquad (j=1,\cdots,n) \tag{2.17.21}$$

を満たすとき**確率行列**(stochastic matrix)，さらに

$$\sum_{j=1}^{n}a_{ij}=1 \qquad (i=1,\cdots,n) \tag{2.17.22}$$

を満たすとき**二重確率行列**(doubly stochastic matrix)と呼ばれる．

van der Waerden(ヴァン・デル・ワルデン)は，1926 年に「n 次の二重確率行列 A のパーマネントは

$$\mathrm{per}\,A\geqq\frac{n!}{n^n} \quad (\sim\sqrt{2\pi n}\,\mathrm{e}^{-n}) \tag{2.17.23}$$

を満たし，等号が成立するのは $A=\frac{1}{n}J_n$ のとき，そしてそのときに限るであろう」という予想を提出した．これが契機となってパーマネントに関する多くの研究がなされたが，1981 年に D. I. Falikman と G. P. Egorychev によって，独立に，この予想が正しいことが示された．

演習問題

2.1 体 K とその上の演算×，＋の代りに，分配束 L とその上の演算∧，∨を採用して，行列の間の演算を

$$C=A\vee B:\quad c_{ij}=a_{ij}\vee b_{ij},$$
$$B=\lambda A\quad(\lambda\in L):\quad b_{ij}=\lambda\wedge a_{ij},$$
$$C=A\cdot B:\quad c_{ij}=\bigvee_k(a_{ik}\wedge b_{kj})$$

で定義したとき，通常の行列算とどこまで類似の性質があるかを確かめよ．特に，零ベクトル，零行列，単位行列に対応するものは何か．

［略解］ 分配束 L の最大元を 1，最小元を 0 とすれば，零ベクトル，零行列，単位行列の定義は，形式的には 2.2 節におけるものと同じ．行列の間の諸演算も，“逆演算”を含まないものは，形式的には同じ性質を有する．

2.2 実数と∞とよりなる集合 $\mathscr{A}=\mathbb{R}\cup\{\infty\}$ の上に二つの演算＋, min が定義されている代数系を考える．（∞と∞，あるいは∞と実数の間の演算＋, min は“常識的”に定義する．）$\mathscr{A}$ の上の行列の間の演算を

$$C=A\min B:\quad c_{ij}=\min(a_{ij},b_{ij}),$$
$$B=\lambda A\quad(\lambda\in\mathscr{A}):\quad b_{ij}=\lambda+a_{ij},$$
$$C=A\cdot B:\quad c_{ij}=\min_k\{a_{ik}+b_{kj}\}$$

と定義して，演習問題 2.1 と同様のことを考察せよ．

n 個の都市 $t_1,\cdots,t_n$ があって t_j を t_i に直接結ぶ道の長さが a_{ij} であるとする（直接結ぶ道がないときには $a_{ij}=\infty$ とする）．このとき行列 $A=[a_{ij}]$ から，代数系 $\mathscr{A}$ 上で

$$A^{(1)}=A,\quad A^{(r+1)}=A\cdot A^{(r)}\qquad(r=1,2,\cdots)$$

により定められる行列 $A^{(r)}=[a_{ij}^{(r)}]$ の要素 $a_{ij}^{(r)}$ は何を表すか．

［略解］ この代数系における演算 min, ＋を体における演算＋, ×に擬えてみよ．すべての要素が∞のベクトル（行列）が零ベクトル（行列），対角要素が 0 で他のすべての要素が∞の行列が単位行列の役割を果たす．$a_{ii}=0$ とおくことにすれば，$a_{ij}^{(r)}$ は「たかだか $r-1$ 個の都市を通過して都市 t_j から都市 t_i に行く道の中で最短なものの長さ」である．

2.3 実数体 $\mathbb{R}$ の上の 2 次正方行列

$$\mathbf{1}=\begin{bmatrix}1&0\\0&1\end{bmatrix},\quad \mathbf{i}=\begin{bmatrix}0&1\\-1&0\end{bmatrix}$$

から行列の定数倍，和，積によって作られる 2 次正方行列の全体は，複素数体 $\mathbb{C}$ と同型であることを示せ．複素数 $\lambda+\mu\mathrm{i}$ に対応する行列 $\begin{bmatrix}\lambda&\mu\\-\mu&\lambda\end{bmatrix}$ の行列式は複素数の絶対値の 2 乗 $|\lambda+\mathrm{i}\mu|^2$ に対応することを確かめよ．

2.4 実数体 $\mathbb{R}$ の上の 4 次正方行列

$$\mathbf{1}=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix},\quad \mathbf{i}=\begin{bmatrix}0&0&1&0\\0&0&0&1\\-1&0&0&0\\0&-1&0&0\end{bmatrix},$$

$$\mathbf{j}=\begin{bmatrix}0&1&0&0\\-1&0&0&0\\0&0&0&-1\\0&0&1&0\end{bmatrix},\quad \mathbf{k}=\begin{bmatrix}0&0&0&-1\\0&0&1&0\\0&-1&0&0\\1&0&0&0\end{bmatrix}$$

から行列の定数倍，和，積によって作られる 4 次正方行列の全体は，**四元数体**(quaternion)と同型であることを示せ．また，上の 4 個の行列を 2 次行列の Kronecker 積の形に表せ．四元数 $\lambda\mathbf{1}+\mu\mathbf{i}+\nu\mathbf{j}+\sigma\mathbf{k}$ に対応する行列

$$\begin{bmatrix}\lambda&\nu&\mu&-\sigma\\-\nu&\lambda&\sigma&\mu\\-\mu&-\sigma&\lambda&-\nu\\\sigma&-\mu&\nu&\lambda\end{bmatrix}$$

の行列式は何を意味するか．

［略解］ 下式の右辺の $\mathbf{i},\mathbf{1}$ は演習問題 2.3 の 2 次正方行列を表すものとすれば，

$$\mathbf{1}=\mathbf{1}\otimes\mathbf{1},\quad \mathbf{i}=\mathbf{i}\otimes\mathbf{1},\quad \mathbf{j}=\begin{bmatrix}1&0\\0&-1\end{bmatrix}\otimes\mathbf{i},\quad \mathbf{k}=\begin{bmatrix}0&-1\\-1&0\end{bmatrix}\otimes\mathbf{i}$$

である．$\lambda\mathbf{1}+\mu\mathbf{i}+\nu\mathbf{j}+\sigma\mathbf{k}$ に対応する行列の行列式は $|\lambda^2+\mu^2+\nu^2+\sigma^2|^2$ に等しい．

2.5 実数体 $\mathbb{R}$ の上の n 次正方行列 $A=[a_{ij}]$ は，

$$a_{ij}\geqq 0\quad (i,j=1,\cdots,n),\quad \sum_{i=1}^{n}a_{ij}=1\quad (j=1,\cdots,n)$$

を満たすとき**確率行列**と呼ばれる(2.17.2 項)．A,B が共に確率行列のとき，積 AB も確率行列であることを示せ．

2.6 2.5.2 項の交代化演算のもつ諸性質を，$n=3$ の場合について具体的に書き下して確認せよ．

2.7 次の行列の行列式の値を計算せよ．

(i) $\begin{bmatrix}8&4&17\\24&10&57\\33&16&72\end{bmatrix}$ (ii) $\begin{bmatrix}2&1&3&4\\1&-3&1&5\\3&1&6&-2\\4&5&-2&-1\end{bmatrix}$ (iii) $\begin{bmatrix}3&-1&-1\\-1&3&-1\\-1&-1&3\end{bmatrix}$

(iv) $n\left[\begin{matrix} a & b & \cdots\cdots & b \\ b & a & b & \vdots \\ \vdots & \ddots & \ddots & b \\ b & \cdots\cdots & b & a \end{matrix}\right]$　(v) $n\left[\begin{matrix} a & b & & & O \\ b & a & b & & \\ & b & \ddots & \ddots & \\ & & \ddots & \ddots & b \\ O & & & b & a \end{matrix}\right]$　(vi) $\begin{bmatrix} 0 & x & y & z \\ -x & 0 & u & v \\ -y & -u & 0 & w \\ -z & -v & -w & 0 \end{bmatrix}$

(vii) $\begin{bmatrix} n & n-1 & n-2 & \cdots & 1 \\ n-1 & n & n-1 & \cdots & 2 \\ \vdots & n-1 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & n-1 \\ 1 & 2 & \cdots & n-1 & n \end{bmatrix}$　(viii) $\begin{bmatrix} 1 & 1/2 & 1/3 & \cdots & 1/n \\ 1/2 & 1 & 2/3 & \cdots & 2/n \\ 1/3 & 2/3 & 1 & \cdots & 3/n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1/n & 2/n & 3/n & \cdots & 1 \end{bmatrix}$

［略解］ (i) -6； (ii) 568； (iii) 16； (iv) 行 $1,\cdots,n-1$ を行 n に加え，次に列 n を列 $1,\cdots,n-1$ から引いてみよ(LU 分解をおこなって法則性を発見することも試みよ)→$(a-b)^{n-1}\ (a+(n-1)b)$；(v) この形の n 次の行列式の値を f_n とすれば $f_n-af_{n-1}+b^2f_{n-2}=0$，$f_1=a$，$f_2=a^2-b^2$，$f_3=a^3-2ab^2,\cdots$であるので，一般の n に対してはこの差分方程式の解として

$$f_n=\frac{1}{\sqrt{a^2-4b^2}}\left[\left(\frac{a+\sqrt{a^2-4b^2}}{2}\right)^{n+1}-\left(\frac{a-\sqrt{a^2-4b^2}}{2}\right)^{n+1}\right]$$

が得られる；(vi) $(xw-yv+zu)^2$；(vii) まず行 $i+1$ から行 i を引き $(i=n-1,\cdots,2,1)$，次に行 i から行 $i+1$ を引く $(i=n-1,\cdots,2,1)$→$2^{n-2}(n+1)$；(viii) $L=[l_{ij}]$ $(l_{ij}=1/i\,(i\geqq j),=0\,(i<j))$ なる下三角行列を用いて，この行列は $L\cdot\mathrm{diag}[1,3,5,\cdots,2n-1]\cdot L^{\mathrm{T}}$ と表せることに注意せよ→$(2n-1)!/(2^{n-1}(n!)^2\ (n-1)!)$．

2.8　A,B が r 次の正方行列のとき

(i)　$\det\begin{bmatrix} A & B \\ B & A \end{bmatrix}=\det(A+B)\cdot\det(A-B)$,

(ii)　$\det\begin{bmatrix} A & B \\ -B & A \end{bmatrix}=\det(A+\mathrm{i}B)\cdot\det(A-\mathrm{i}B)$,

(iii)　$\det\begin{bmatrix} A & I_r \\ -I_r & B \end{bmatrix}=\det(I_r+AB)$

であることを示せ．

［略解］ (i) 行 $1,\cdots,r$ を行 $r+1,\cdots,2r$ にそれぞれ加え，次に列 $1,\cdots,r$ から列 $r+1,\cdots,2r$ をそれぞれ引く．(ii) (i) にならえ．(iii) $\begin{bmatrix} I_r & A \\ O & I_r \end{bmatrix}\cdot\begin{bmatrix} A & I_r \\ -I_r & B \end{bmatrix}=\begin{bmatrix} O & I_r+AB \\ -I_r & AB \end{bmatrix}$ の両辺の行列式を取れ．

2.9　$(A\cdot(A+I))^{-1}=A^{-1}-(A+I)^{-1}$

［略解と注］ A と I とが交換可能であることに注意して両辺に $A\cdot(A+I)$ を掛けてみればよい．実質的には $\dfrac{1}{x(x+1)}=\dfrac{1}{x}-\dfrac{1}{x+1}$ の行列版である．

2.10　Pfaffian に関する以下の等式が成立することを確かめよ．ここで，A, B はそれぞれ $2m$次，$2n$次の反対称行列である．

(i)　$\mathrm{Pf}\begin{bmatrix} A & O \\ O & B \end{bmatrix} = (\mathrm{Pf}\,A)\cdot(\mathrm{Pf}\,B)$

(ii)　$m=n$ のとき　$\mathrm{Pf}(BAB^{\mathsf{T}}) = (\det B)\cdot(\mathrm{Pf}\,A)$

(iii)　$\mathrm{Pf}(\lambda A) = \lambda^m\,\mathrm{Pf}\,A$

(iv)　$\mathrm{Pf}\,A^{\mathsf{T}} = (-1)^m\,\mathrm{Pf}\,A$

［略解］　(ii)以外はほとんど自明．(ii)については補章 A.5 節も参照のこと．

2.11　すべての n 次正方行列と交換可能な行列は，“スカラー行列” $\lambda I_n\,(\lambda \in K)$ に限ることを示せ．

［略解］　(i,j)要素が 1 で，他の要素がすべて 0 の n 次正方行列 $E(i,j)$ との積を考えよ．

2.12　問題 2.7 の行列の LU 分解，逆行列を求めよ．

2.13　整数行列($K=\mathbb{Z}$ の上の行列)

$$A=\begin{bmatrix} 6 & 8 & 12 \\ 3 & 7 & 21 \\ 5 & 11 & 31 \end{bmatrix}$$

の単因子および行列式因子を定めよ．

［略解］　単因子 $e_1(A)=1$，$e_2(A)=2$，$e_3(A)=6$；行列式因子 $d_1(A)=1$，$d_2(A)=2$，$d_3(A)=12$．

2.14　多項式行列($K=\mathbb{R}[\lambda]$の上の行列)

$$A=\begin{bmatrix} \lambda^2+\lambda-1 & \lambda^3+\lambda^2 & 2\lambda^2-1 \\ \lambda^2+\lambda & \lambda^3+\lambda^2+\lambda & 2\lambda^2 \\ 2\lambda^2+\lambda-1 & 2\lambda^3+\lambda^2 & 4\lambda^2-\lambda-1 \end{bmatrix}$$

の単因子および行列式因子を定めよ．

［略解］　単因子 $e_1(A)=1$，$e_2(A)=\lambda$，$e_3(A)=\lambda(\lambda-1)$；行列式因子 $d_1(A)=1$，$d_2(A)=\lambda$，$d_3(A)=\lambda^2(\lambda-1)$．

2.15　$K=\mathbb{C}$ のとき正方行列 A は以下の各々の 3 条件のうちの 2 条件を満たせば残りの 1 条件も満たすことを確認せよ．

〈a〉　実行列である；直交行列である；ユニタリ行列である．

〈b〉　対称行列である；直交行列である；対合行列($A^2=I$)である．

〈c〉　Hermite 行列である；ユニタリ行列である；対合行列である．

3
ベクトル空間

線形代数においてベクトル空間の概念は最も基本的であるが，有限次元のベクトル空間の性質を調べるための技術的な準備は既に第2章でほとんど済んでいるので，本章の話は超特急で進めることができる．

3.1　ベクトル空間

3.1.1　ベクトル空間の定義

(空でない)集合 V が**体 K の上のベクトル空間**(vector space over a field K)であるとは，V の任意の元 $\boldsymbol{x},\boldsymbol{y}$ に対して**和** $\boldsymbol{x}+\boldsymbol{y}$ が V の元として一意に定まり，V の任意の元 $\boldsymbol{x}$ と K の任意の元 λ とに対して $\boldsymbol{x}$ の λ 倍(**スカラー倍**という) $\lambda\cdot\boldsymbol{x}$ が V の元として一意に定まり，これらの演算が以下の条件(i)～(viii)を満たすことである．このとき，V の元を**ベクトル**(vector)，K の元を**スカラー**(scalar)と呼ぶ．(誤解の恐れのない限り積演算記号・を省略することも多い．)

(i)　$\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\in V$ に対して $(\boldsymbol{x}+\boldsymbol{y})+\boldsymbol{z}=\boldsymbol{x}+(\boldsymbol{y}+\boldsymbol{z})$．　[和の結合則]

(ii)　$\boldsymbol{x},\boldsymbol{y}\in V$ に対して $\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{y}+\boldsymbol{x}$．　[和の交換則]

(iii)　すべてのベクトル $\boldsymbol{x}\in V$ に対して $\boldsymbol{x}+\boldsymbol{0}=\boldsymbol{x}$ となるような $\boldsymbol{0}(\in V)$ [**零ベクトル**(null vector)と呼ぶ]が存在する．

(iv)　$\boldsymbol{x}\in V$ に対して $\boldsymbol{x}+(-\boldsymbol{x})=\boldsymbol{0}$ となるようなベクトル $-\boldsymbol{x}(\in V)$ が存在する．　[和に関する $\boldsymbol{x}$ の逆元]

(v)　K の乗法単位元 1 とベクトル $\boldsymbol{x}\in V$ に対して $1\cdot\boldsymbol{x}=\boldsymbol{x}$．

(vi)　$\lambda,\mu\in K$，$\boldsymbol{x}\in V$ に対して $(\lambda\mu)\cdot\boldsymbol{x}=\lambda\cdot(\mu\cdot\boldsymbol{x})$．
[スカラー倍の結合則]

(vii)　$\lambda\in K$，$\boldsymbol{x},\boldsymbol{y}\in V$ に対して $\lambda\cdot(\boldsymbol{x}+\boldsymbol{y})=\lambda\cdot\boldsymbol{x}+\lambda\cdot\boldsymbol{y}$．

[ベクトルの分配則]

(viii) $\lambda,\mu\in K$, $\boldsymbol{x}\in V$ に対して $(\lambda+\mu)\cdot\boldsymbol{x}=\lambda\cdot\boldsymbol{x}+\mu\cdot\boldsymbol{x}$.

[スカラーの分配則]

注意 3.1.1 零ベクトルはただ一つしかない($\boldsymbol{0}$ と $\boldsymbol{0}'$ が零ベクトルであったとすると，(ii)，(iii)により $\boldsymbol{0}+\boldsymbol{0}'=\boldsymbol{0}'+\boldsymbol{0}=\boldsymbol{0}=\boldsymbol{0}'$)．$\boldsymbol{x}\in V$ に対して $-\boldsymbol{x}$ はただ一つしかない($\boldsymbol{x}+\boldsymbol{y}=\boldsymbol{0}$ であるとすると，$(-\boldsymbol{x})+(\boldsymbol{x}+\boldsymbol{y})=(-\boldsymbol{x})+\boldsymbol{0}=-\boldsymbol{x}$；この左辺は(i)，(ii)により，$=\boldsymbol{0}+\boldsymbol{y}=\boldsymbol{y}$；すなわち $\boldsymbol{y}=-\boldsymbol{x}$)．$\boldsymbol{y}+(-\boldsymbol{x})$ のことを $\boldsymbol{y}-\boldsymbol{x}$ とも記す．$\boldsymbol{x},\boldsymbol{y}\in V$ に対して $\boldsymbol{x}+\boldsymbol{z}=\boldsymbol{y}$ を満たす $\boldsymbol{z}\in V$ がただ一つ存在する($\boldsymbol{z}=\boldsymbol{y}-\boldsymbol{x}$ である)．

注意 3.1.2 (v)，(viii)より，$(1+0)\cdot\boldsymbol{x}=\boldsymbol{x}+0\cdot\boldsymbol{x}=1\cdot\boldsymbol{x}=\boldsymbol{x}$；そこで $\boldsymbol{0}=0\cdot\boldsymbol{x}$ である．さらに，$\boldsymbol{0}=0\cdot\boldsymbol{x}=(1-1)\cdot\boldsymbol{x}=\boldsymbol{x}+(-1)\cdot\boldsymbol{x}$ より，$-\boldsymbol{x}=(-1)\cdot\boldsymbol{x}$ である．また，(vi)により，$\lambda\in K$，$\boldsymbol{x}\in V$ に対して，$\lambda\cdot\boldsymbol{0}=\lambda\cdot(0\cdot\boldsymbol{x})=(\lambda\cdot 0)\cdot\boldsymbol{x}=0\cdot\boldsymbol{x}=\boldsymbol{0}$.

例 3.1.1 数学自身の諸分野，応用諸分野において，ベクトル空間とみなすことのできる(そしてそうすることが役に立つ)対象は多数ある．その一部については演習問題 3.1 を参照． □

3.1.2 ベクトルの独立性，従属性

体 K の上のベクトル空間 V の m 本のベクトルの(有限)集合 $X=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_m\}$ のスカラーの集合 $\{\lambda_1,\cdots,\lambda_m\}$ $(\lambda_i\in K)$ を係数とする**線形結合**(linear combination；**一次結合**ともいう)とは，ベクトル

$$\lambda_1\boldsymbol{x}_1+\cdots+\lambda_m\boldsymbol{x}_m$$

のことである．X が**線形従属**(linearly dependent；**一次従属**ともいう)であるとは，すべてが 0 ではないスカラー $\{\lambda_1,\cdots,\lambda_m\}$ を係数とする線形結合で $\boldsymbol{0}$ となるものがあること，すなわち

$$\lambda_1\boldsymbol{x}_1+\cdots+\lambda_m\boldsymbol{x}_m=\boldsymbol{0} \tag{3.1.1}$$

であることをいう．線形従属でないとき**線形独立**(linearly independent；**一次独立**)であるという．すなわち $X=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m\}$ $(\boldsymbol{x}_i\in V)$ が(3.1.1)を満たすのが $\lambda_1=\lambda_2=\cdots=\lambda_m=0$ のときに限るとき，X は線形独立であるという．$X=\emptyset$ (空集合)は線形独立であると定義する．

注意 3.1.3 $X=\{\boldsymbol{x}_1\}$ が線形従属であるというのは $\lambda_1\boldsymbol{x}_1=\boldsymbol{0},\lambda_1\neq 0$ ということで

あるから，$\boldsymbol{x}_1=(1/\lambda_1)\cdot\boldsymbol{0}=\boldsymbol{0}$，すなわち $\boldsymbol{x}_1$ が零ベクトルであるということである．$X=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m\}$が線形独立で $X\cup\{\boldsymbol{y}\}$が線形従属ならば，

$$\lambda_1\boldsymbol{x}_1+\cdots+\lambda_m\boldsymbol{x}_m+\mu\boldsymbol{y}=\boldsymbol{0} \tag{3.1.2}$$

となる $\lambda_1,\cdots,\lambda_m,\mu$($\lambda_1=\cdots=\lambda_m=\mu=0$ ではない)が存在するが，このとき $\mu\neq0$ である($\mu=0$ とすると，$\lambda_1,\cdots,\lambda_m$ の中に 0 でないものがなければならないことになり X の独立性に反する)．そこで(3.1.2)の両辺に $1/\mu$ を掛けて，

$$\boldsymbol{y}=\left(-\frac{\lambda_1}{\mu}\right)\boldsymbol{x}_1+\left(-\frac{\lambda_2}{\mu}\right)\boldsymbol{x}_2+\cdots+\left(-\frac{\lambda_m}{\mu}\right)\boldsymbol{x}_m. \tag{3.1.3}$$

すなわち $\boldsymbol{y}$ は X の線形結合で表される(3.1.1 項，(vii)および注意 3.1.1 も参照)．

さらに，この表現は一意的である．すなわち

$$\boldsymbol{y}=\lambda_1'\boldsymbol{x}_1+\cdots+\lambda_m'\boldsymbol{x}_m=\lambda_1''\boldsymbol{x}_1+\cdots+\lambda_m''\boldsymbol{x}_m \tag{3.1.4}$$

と表されたとすると，

$$(\lambda_1'-\lambda_1'')\boldsymbol{x}_1+\cdots+(\lambda_m'-\lambda_m'')\boldsymbol{x}_m=\boldsymbol{0} \tag{3.1.5}$$

となるが，X が線形独立であるから $\lambda_1'=\lambda_1'',\cdots,\lambda_m'=\lambda_m''$ でなければならない．

線形独立・従属の関係は次の基本的な性質(1)，(2)，(3) をもつ．

(1) $\emptyset$ は線形独立である．

(2) $X=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m\}$が線形独立ならその任意の部分集合 $Y(\subset X)$も線形独立である．

(3) $X=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$, $Y=\{\boldsymbol{y}_1,\cdots,\boldsymbol{y}_m\}$が線形独立で $|X|<|Y|$(すなわち $n<m$)ならば，$X\cup\{\boldsymbol{y}_{j_0}\}$も線形独立であるような $Y-X$ の元 $\boldsymbol{y}_{j_0}$ が存在する．

略証 (1)は定義．(2)：一般性を失うことなく $Y=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_l\}$($1\leqq l<m$)とする．$Y$ が線形従属なら $\lambda_1\boldsymbol{x}_1+\cdots+\lambda_l\boldsymbol{x}_l=\boldsymbol{0}$ となるようなすべてが 0 ではない $\lambda_1,\cdots,\lambda_l(\in K)$がある．このとき $\lambda_{l+1}=\cdots=\lambda_m=0$ とおけば $\lambda_1\boldsymbol{x}_1+\cdots+\lambda_m\boldsymbol{x}_m=\boldsymbol{0}$，すなわち X も線形従属となる．(3)：すべての i について $X\cup\{\boldsymbol{y}_i\}$が線形従属であるとすると，注意 3.1.3 により，適当な係数 a_{ij} を用いて $\boldsymbol{y}_i$ は X の線形結合

$$\boldsymbol{y}_i=\sum_{j=1}^{n}a_{ij}\boldsymbol{x}_j \qquad (i=1,\cdots,m) \tag{3.1.6}$$

を用いて表される．このとき，$m\times n$ 行列 $A=[a_{ij}]$は，$m>n\geqq\mathrm{rank}\,A$ であるから，左からある m 次正方正則行列 S を掛けることによって，SA の最後の $m-\mathrm{rank}\,A(\geqq1)$個の行を 0 にすることができる(2.15.1 項，基本変形を参照)．S の第 m 行を $s_{mi}(i=1,\cdots,m)$とすれば，S の正則性から $s_{m1},\cdots,s_{mm}$ の少なく

とも1個は非零であり，

$$\sum_{i=1}^{m} s_{mi}\boldsymbol{y}_i=\sum_{j=1}^{n}\left(\sum_{i=1}^{m} s_{mi}a_{ij}\right)\boldsymbol{x}_j=\boldsymbol{0} \tag{3.1.7}$$

となるが，これは Y の独立性に反する．■

注意 3.1.4　ある集合 E の各有限部分集合に対して，上記の(1)，(2)，(3)が満たされるように“独立”であるか“従属”であるかが定められており，独立な部分集合の元の数が有界であるとき，E のことを**マトロイド**(matroid)という．マトロイドの理論は，その応用可能性の広がりとも相俟って，近年大いに発展している．

線形独立・従属性に関する多くの関係は，このより一般的なマトロイドについても成り立つものである．たとえば，(3)から

・極大な独立集合の要素数はみな等しい

ことがわかる(独立集合 X が極大であるとは，任意の $\boldsymbol{y}\notin X$ に対して $X\cup\{\boldsymbol{y}\}$ が従属であることである)．演習問題3.2も見よ．

3.1.3　ベクトル空間の次元と基底

ベクトル空間 V の極大線形独立集合の要素数のことを V の**次元**(dimension)といい，$\dim V$ と記す．(注意3.1.4により $\dim V$ は極大線形独立集合の選び方によらず確定する)．$\dim V=n$ のベクトル空間を V_n のように記すことが多い．

$\dim V$ は無限大でもありうるが，以下ではそれが有限である場合に話を限る．このとき，V の極大線形独立集合のことを V の**基底**(basis)と呼ぶ．零ベクトルのみよりなるベクトル空間 $V=\{\boldsymbol{0}\}$ の次元は0，基底は $\emptyset$ であるとする．また，3.1.2項の線形独立性の性質(3)によれば，任意の線形独立集合 X に対してそれを含む基底が存在する．

$B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ が V の基底であれば，注意3.1.3により，任意のベクトル $\boldsymbol{x}$ $(\in V)$ は B の線形結合で一意的に表される：

$$\boldsymbol{x}=\sum_{j=1}^{n} x_j\boldsymbol{e}_j. \tag{3.1.8}$$

縦ベクトル $[x_1,\cdots,x_n]^{\top}$ (あるいはその成分 $x_1,\cdots,x_n$)を，基底 B に関するベクトル $\boldsymbol{x}$ の**座標**(成分)(coordinates)ということがある．(3.1.8)は記号的に

$$\boldsymbol{x}=[\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \tag{3.1.9}$$

とも書ける．

V_n には一般にいくつもの基底がある．$B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}, B'=\{\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n'\}$が二つの基底であるとき，$B'$ のベクトルを B の線形結合で表す式と B のベクトルを B' の線形結合で表す式を

$$\boldsymbol{e}_j'=\sum_{i=1}^{n} s_{ij}\boldsymbol{e}_i \qquad (j=1,\cdots,n), \tag{3.1.10a}$$

$$\boldsymbol{e}_j=\sum_{i=1}^{n} t_{ij}\boldsymbol{e}_i' \qquad (j=1,\cdots,n) \tag{3.1.10b}$$

とする．(3.1.10a)を(3.1.10b)に代入すると

$$\boldsymbol{e}_j=\sum_{i=1}^{n}\left(\sum_{k=1}^{n} s_{ik}t_{kj}\right)\boldsymbol{e}_i \qquad (j=1,\cdots,n). \tag{3.1.11}$$

このことは，n 次正方行列 $S=[s_{ij}]$と $T=[t_{ij}]$とが互いに逆行列である，すなわち $ST=I_n$ であることを表している((3.1.10b)を(3.1.10a)に代入すれば $TS=I_n$ が得られるが，逆行列の性質(注意 2.10.1)により，結論は同じである)．同一のベクトル $\boldsymbol{x}$ の座標成分については

$$\boldsymbol{x}=\sum_{j=1}^{n} x_j'\boldsymbol{e}_j'=\sum_{j=1}^{n} x_j'\sum_{i=1}^{n} s_{ij}\boldsymbol{e}_i=\sum_{i=1}^{n}\left(\sum_{j=1}^{n} s_{ij}x_j'\right)\boldsymbol{e}_i \tag{3.1.12}$$

より，B に関する座標成分を B' に関するもので表す式

$$x_i=\sum_{j=1}^{n} s_{ij}x_j' \qquad (i=1,\cdots,n) \tag{3.1.13a}$$

および B' に関する成分を B に関するもので表す式

$$x_i'=\sum_{j=1}^{n} t_{ij}x_j \qquad (i=1,\cdots,n) \tag{3.1.13b}$$

が得られる．(3.1.9)のような表し方をすると，(3.1.10)，(3.1.13)は形式的に

$$[\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n']=[\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]S, \qquad [\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]=[\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n']T, \tag{3.1.10$'$}$$

$$\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}=S\begin{bmatrix}x_1'\\ \vdots\\ x_n'\end{bmatrix}, \qquad \begin{bmatrix}x_1'\\ \vdots\\ x_n'\end{bmatrix}=T\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \tag{3.1.13$'$}$$

と書ける．(3.1.10)の基底の間の変換と(3.1.13)の座標成分の間の変換とは，互

いに**反傾**(contragredient)な関係にあるという(→3.4.3項).

$B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$が V_n の一つの基底であるとき，任意の n 次正則行列 S によって(3.1.10)，(3.1.10′)で表される $B'=\{\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n'\}$がやはり V_n の基底となることも明らかであろう.

基底 B を固定すれば，ベクトル空間 V_n のベクトル $\boldsymbol{x}$ はその座標成分からなる n 次元縦ベクトル(2.1節)と一対一に対応し，その対応関係のもとで3.1.1項のベクトルの間の諸演算などは2.2節の縦ベクトルの間の諸演算に対応する．そこで，有限次元ベクトル空間 V に一つの基底を選んで固定し，V のベクトルをその座標成分の作る縦ベクトルと同一視して論じることが多い．このとき，(3.1.9)を，基底を省略して，

$$\boldsymbol{x}=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \tag{3.1.9′}$$

のように書く．このとき，基底ベクトル $\boldsymbol{e}_i$ は第 i 成分のみが1で他の成分がみな0の縦ベクトル $\boldsymbol{e}_i=[\overset{1}{0},\overset{\cdots\cdots}{\cdots},\overset{i-1}{0},\overset{i}{1},\overset{i+1}{0},\overset{\cdots\cdots}{\cdots},\overset{n}{0}]^{\mathsf{T}}$ と表される.

3.1.4 行列の列の独立性と行列の階数

$m\times n$ 型の行列 $A=[a_{ij}]$の n 本の列ベクトルを，V_m の n 個のベクトル $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ のある基底 B に関する座標成分を表す縦ベクトルとみると，3.1.2項において述べたように，m 次正則行列 S を A の左から掛けた $A'=SA$ の列ベクトルは別の基底に関する $\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$ の座標成分を表す．$\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$の(部分集合の)線形独立・従属性は，したがって，$S$ を適当に選んで，A' の列ベクトルの線形独立・従属性の形にして調べることができる.

rank $A=r\,(r\leqq\min(m,n))$のときには，2.15.1項の基本変形によって

$$A'=\begin{array}{r} \\ r\left\{\vphantom{\begin{matrix}1\\ \ddots\\ 1\end{matrix}}\right. \\ m-r\{ \end{array}\!\!\begin{array}{cc} \overbrace{\hphantom{1\;\;\ddots\;\;O}}^{r} & \overbrace{\hphantom{CC}}^{n-r} \\ \begin{matrix}1 & & O\\ & \ddots & \\ O & & 1\end{matrix} & C \\ \multicolumn{2}{c}{O} \end{array} \tag{3.1.14}$$

という形にできる(列ベクトルの順序は適当に入れ換えられているとする)．この形にしてみれば，A' の左側の r 本の列は線形独立で，右側の $n-r$ 本の列が左

側の r 本の列の線形結合で表せること $\left(\boldsymbol{x}_j=-\sum_{i=1}^{r} c_{ij}\boldsymbol{x}_i\,(j=r+1,\cdots,n)\right.$，$c_{ij}$ は C の (i,j) 要素）は明らかである．したがって，$r<n$ ならば $X=\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n\}$ は線形従属である．$r=n$ ならば X は線形独立である．

このようにベクトルの線形独立・従属性は行列の階数によって判定できる．

3.1.5 ベクトルのノルム

ここでは $K=\mathbb{R}$ または $\mathbb{C}$ とする．

ベクトルの“大きさ”あるいは“長さ”のようなものを表したいときに用いられる概念が**ノルム**(norm)である．第2章で既に何カ所かで**2乗ノルム**を用いた．それは，ベクトル $\boldsymbol{x}$ のある基底 B に関する座標成分 $[x_1,\cdots,x_n]^{\top}$ を用いて

$$\|\boldsymbol{x}\|_2=\left(\sum_{i=1}^{n}|x_i|^2\right)^{1/2} \tag{3.1.15}$$

と定義されるものであった．基底 B と B' を結ぶ基底変換の行列が直交行列（$K=\mathbb{C}$ のときはユニタリ行列）$Q=[q_{ij}]$ であるときには，B' に関する座標成分を $[x_1',\cdots,x_n']^{\top}$ とすると，

$$\begin{aligned}\sum_{i=1}^{n}|x_i'|^2&=\sum_{i=1}^{n}\left|\sum_{j=1}^{n}q_{ij}x_j\right|^2=\sum_{j=1}^{n}\sum_{k=1}^{n}\left(\sum_{i=1}^{n}\bar{q}_{ij}q_{ik}\right)\bar{x}_jx_k\\&=\sum_{j=1}^{n}|x_j|^2\end{aligned} \tag{3.1.16}$$

であるから（2.14.1 項の(2.14.4)および 2.14.2 項の(2.14.24)を見よ），$\|\boldsymbol{x}\|_2$ は基底の直交変換（ユニタリ変換）に関して不変である．

2乗ノルムは次の(1)，(2)，(3)を満たす．

(1)　$\|\boldsymbol{x}\|\geqq 0$；$\|\boldsymbol{x}\|=0 \Longleftrightarrow \boldsymbol{x}=\boldsymbol{0}$.

(2)　$\|\lambda\boldsymbol{x}\|=|\lambda|\cdot\|\boldsymbol{x}\|$　　$(\lambda\in K,\boldsymbol{x}\in V_n)$.

(3)　$\|\boldsymbol{x}+\boldsymbol{y}\|\leqq\|\boldsymbol{x}\|+\|\boldsymbol{y}\|$　　$(\boldsymbol{x},\boldsymbol{y}\in V_n)$.

((1)，(2)は定義から明らか．(3)は**三角不等式**と呼ばれるが，これが成り立つことは以下でより一般的な場合について示す．)

ベクトル空間 V から実数 $\mathbb{R}$ への関数で上記の性質(1)，(2)，(3)を満たすものを，一般に，**ノルム**(norm)と呼ぶ．ノルムにはいろいろなものがあるが，任意の二つのノルム $\|\cdot\|$ と $\|\cdot\|'$ に対して正定数 a,b が存在して

$$a\|\boldsymbol{x}\|\leqq\|\boldsymbol{x}\|'\leqq b\|\boldsymbol{x}\| \qquad (\forall\,\boldsymbol{x}\in V) \tag{3.1.17}$$

が成り立つという意味で，すべてのノルムは互いに等価である(→演習問題 3.4)．

例 3.1.2 ノルムとして使い勝手のよいものは，すでに例示した 2 乗ノルムの他に

(N 1) **絶対値和ノルム**： $\|\boldsymbol{x}\|_1=\sum_{i=1}^{n}|x_i|,$ (3.1.18)

(N∞) **最大値ノルム**： $\|\boldsymbol{x}\|_\infty=\max_{i=1,\cdots,n}|x_i|$ (3.1.19)

がある．これらはもっと広い

(Np) **p 乗ノルム**： $\|\boldsymbol{x}\|_p=\left(\sum_{i=1}^{n}|x_i|^p\right)^{1/p}$ (3.1.20)

というノルムの族(ただし $p\geqq1$)の特殊なものである．(2 乗ノルムはもちろん $p=2$ の場合，絶対値和ノルムは $p=1$ の場合であり，最大値ノルムは $p=\infty$ に相当するとみなせる．)

絶対値和ノルム，最大値ノルムがノルムの 3 条件(1)，(2)，(3)を満たすこと，そして p 乗ノルムが(1)，(2)を満たすことは明らかであるが，$1<p<\infty$ のとき p 乗ノルムが(3)を満たすことは，次のようにして示される．まず，関数 $f(t)=t^p$ が $t>0$ において凸である($f''(t)>0$)ことから，任意の $\lambda\in(0,1)$ と任意の非負実数 ξ,η に対して

$$f(\lambda\xi+(1-\lambda)\eta)\leqq\lambda f(\xi)+(1-\lambda)f(\eta) \tag{3.1.21}$$

であることに注意する．$\boldsymbol{x},\boldsymbol{y}$ の少なくとも一方が零ベクトルの場合(3)が成り立つことは明らかであるから，$\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{y}\neq\boldsymbol{0}$ として，

$$\lambda=\frac{\|\boldsymbol{x}\|_p}{\|\boldsymbol{x}\|_p+\|\boldsymbol{y}\|_p},\qquad 1-\lambda=\frac{\|\boldsymbol{y}\|_p}{\|\boldsymbol{x}\|_p+\|\boldsymbol{y}\|_p} \tag{3.1.22}$$

そして

$$\xi_i=\frac{x_i}{\|\boldsymbol{x}\|_p},\qquad \eta_i=\frac{y_i}{\|\boldsymbol{y}\|_p} \tag{3.1.23}$$

とおく．すると，

$$\begin{gathered}\sum_{i=1}^{n}|\xi_i|^p=1,\quad \sum_{i=1}^{n}|\eta_i|^p=1,\\ \sum_{i=1}^{n}|\lambda\xi_i+(1-\lambda)\eta_i|^p=\left(\frac{\|\boldsymbol{x}+\boldsymbol{y}\|_p}{\|\boldsymbol{x}\|_p+\|\boldsymbol{y}\|_p}\right)^p\end{gathered} \tag{3.1.24}$$

となるので，

$$\sum_{i=1}^{n}|\lambda\xi_i+(1-\lambda)\eta_i|^p \leqq \sum_{i=1}^{n}(\lambda|\xi_i|+(1-\lambda)|\eta_i|)^p$$
$$\leqq \sum_{i=1}^{n}(\lambda|\xi_i|^p+(1-\lambda)|\eta_i|^p)$$
$$=\lambda\sum_{i=1}^{n}|\xi_i|^p+(1-\lambda)\sum_{i=1}^{n}|\eta_i|^p=\lambda+(1-\lambda)=1 \tag{3.1.25}$$

が得られる．これは三角不等式(3)に他ならない．□

3.2　部分空間，補空間

3.2.1　部分空間

ベクトル空間 V の(空でない)部分集合 U が(V における と同じ演算に関して)それ自身一つのベクトル空間になっているとき，U を V の**部分空間**(subspace)であるという．

注意 3.2.1　代数学の言葉でいうと3.1.1項の(i)～(viii)は，ベクトル空間が K-加群であることを表している．部分空間は，代数学の言葉でいえば，部分代数系のことである．V がベクトル空間であるから，その部分集合 U が部分空間であるためには次の2条件を満たせばよい：

(イ)　$\boldsymbol{x}, \boldsymbol{y} \in U$ ならば $\boldsymbol{x}+\boldsymbol{y} \in U$,

(ロ)　$\boldsymbol{x} \in U$, $\lambda \in K$ ならば $\lambda\boldsymbol{x} \in U$.

((イ)，(ロ)は U の中で和とスカラー倍が定義されているということであるから，U がベクトル空間であるためには当然満たされなければならないが，これだけで十分である．すなわち，ベクトル空間の条件(i)～(viii)のうち，(i)，(ii)，(v)，(vi)，(vii)，(viii)は，V が満たしているから U も満たす．(ロ)で $\lambda=0$ とおけば $0\cdot\boldsymbol{x}=\boldsymbol{0} \in U$，すなわち(iii)が満たされる．また，(ロ)で $\lambda=-1$ とおけば $(-1)\cdot\boldsymbol{x}=-\boldsymbol{x} \in U$，すなわち(iv)が満たされる．)

例 3.2.1　ベクトル空間 V 自身，および零ベクトルのみよりなる部分集合 $\{\boldsymbol{0}\}$ はいずれも V の部分空間である．これらは“自明な”部分空間である．□

注意 3.2.2　U が V の部分空間，W が U の部分空間なら W は V の部分空間である．$\{\boldsymbol{0}\}$ はすべての部分空間の部分空間である．

V の部分空間 U の基底の一つを B_0 とすると，3.1.2項の(3)により，B_0 に V

$-U$ のいくつかのベクトルを追加して $B_0 \cup B_1$ という V の基底を作ることができる．そこで，U が V の部分空間なら，$\dim U \leqq \dim V$ であり，$\dim U = \dim V$ なら $U = V$ である．

ベクトル空間 V のベクトルの集合 $X = \{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m\}$ に対して，X の線形結合で表されるベクトルの全体を

$$L(X) = \{\boldsymbol{y} \in V \mid \boldsymbol{y} = \lambda_1 \boldsymbol{x}_1 + \cdots + \lambda_m \boldsymbol{x}_m \;;\; \lambda_1, \cdots, \lambda_m \in K\} \tag{3.2.1}$$

と記すが，これは，容易に確認できるように，V の部分空間である．これを，X により**生成される部分空間**(subspace generated by X)，あるいは X が**張る部分空間** (subspace spanned by X)という．Y が X の極大線形独立部分集合ならば，Y が $L(X)$ の基底をなすことも明らかであろう．

注意 3.2.3 ベクトル空間はその基底により生成される．

3.2.2 部分空間の共通部分と和

U, U' がベクトル空間 V の部分空間であるとき，U と U' の集合としての共通部分 $U \cap U'$ は V の一つの部分空間をなす(ほとんど自明)．これをこれらの部分空間の**共通部分**(intersection)という．3 個以上の部分空間の共通部分についても同様．

U, U' がベクトル空間 V の部分空間であるとき，U と U' との**和**(sum) $U + U'$ は

$$U + U' = \{\boldsymbol{y} \in V \mid \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{x}' \;;\; \boldsymbol{x} \in U, \boldsymbol{x}' \in U'\} \tag{3.2.2}$$

で定義される．$U + U'$ が V の部分空間であることも明らかである．3 個以上の部分空間の和についても同様．

注意 3.2.4 ベクトル空間の和 $U + U'$ は集合としての和集合(合併集合) $U \cup U'$ とは異なる．もちろん $U \cup U' \subseteqq U + U'$，$U + U' = L(U \cup U')$ ではある．

部分空間 U, U' の共通部分の次元 $\dim(U \cap U')$ と和の次元 $\dim(U + U')$ については次の関係がある．

$$\dim(U + U') + \dim(U \cap U') = \dim U + \dim U' \tag{3.2.3}$$

略証 $U \cap U'$ の基底を $B_0 = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r\}$ とする．U の基底として $B_0 \cup B_1 =$

$\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r, \boldsymbol{b}_1, \cdots, \boldsymbol{b}_s\}$という形のものが選べる(注意 3.2.2)．$U'$ についても $B_0 \cup B_2 = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r, \boldsymbol{c}_1, \cdots, \boldsymbol{c}_t\}$の形の基底を選ぶ．$B_0 \cup B_1 \cup B_2 = \{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_r, \boldsymbol{b}_1, \cdots, \boldsymbol{b}_s, \boldsymbol{c}_1, \cdots, \boldsymbol{c}_t\}$が $U+U'$ を張ることは明らかであるが，これが実は $U+U'$ の基底にもなっている．[それには $B_0 \cup B_1 \cup B_2$ が線形独立であることを示せばよい．実際，$\lambda_i, \mu_j, \nu_k \in K$ に対して

$$\sum_{i=1}^{r} \lambda_i \boldsymbol{a}_i + \sum_{j=1}^{s} \mu_j \boldsymbol{b}_j + \sum_{k=1}^{t} \nu_k \boldsymbol{c}_k = \boldsymbol{0} \tag{3.2.4}$$

となったとすると，

$$\sum \lambda_i \boldsymbol{a}_i + \sum \mu_j \boldsymbol{b}_j = -\sum \nu_k \boldsymbol{c}_k \tag{3.2.5}$$

であるが，(3.2.5)の左辺 $\in U$，右辺 $\in U'$ であるから両辺とも $\in U \cap U'$．したがって，ある $\sigma_i \in K$ を用いて

$$\sum_{i=1}^{r} \sigma_i \boldsymbol{a}_i = -\sum_{k=1}^{t} \nu_k \boldsymbol{c}_k \quad \text{すなわち} \quad \sum_{i=1}^{r} \sigma_i \boldsymbol{a}_i + \sum_{k=1}^{t} \nu_k \boldsymbol{c}_k = \boldsymbol{0} \tag{3.2.6}$$

とならなければならない．$B_0 \cup B_2$ が線形独立であるから $\sigma_i = 0\ (i=1, \cdots, r)$, $\nu_k = 0\ (k=1, \cdots, t)$．すると，(3.2.4)より

$$\sum_{i=1}^{r} \lambda_i \boldsymbol{a}_i + \sum_{j=1}^{s} \mu_j \boldsymbol{b}_j = \boldsymbol{0} \tag{3.2.7}$$

が得られるが，$B_0 \cup B_1$ が線形独立であるから $\lambda_i = 0\ (i=1, \cdots, r)$, $\mu_j = 0\ (j=1, \cdots, s)$．]

$$\begin{aligned} &\dim U = r+s, && \dim U' = r+t, \\ &\dim(U \cap U') = r, && \dim(U+U') = r+s+t \end{aligned} \tag{3.2.8}$$

であるから，(3.2.3)が成り立つ．　∎

部分空間 $U_1, \cdots, U_k$ の和 $U_1 + \cdots + U_k$ は，$\boldsymbol{x}_1 + \cdots + \boldsymbol{x}_k\ (\boldsymbol{x}_i \in U_i)$という形のベクトルの集まりであるが，一般にはこの表し方は一意的でない．すなわち，必ずしも $\boldsymbol{x}_i = \boldsymbol{y}_i$ でなくても

$$\boldsymbol{x}_1 + \cdots + \boldsymbol{x}_k = \boldsymbol{y}_1 + \cdots + \boldsymbol{y}_k \qquad (\boldsymbol{x}_i, \boldsymbol{y}_i \in U_i) \tag{3.2.9}$$

となることがある．特に，(3.2.9)が成り立つのが $\boldsymbol{x}_i = \boldsymbol{y}_i\ (i=1, \cdots, k)$のときに限るとき，すなわち，和 $U_1 + \cdots + U_k$ のベクトルを各部分空間 U_i のベクトル $\boldsymbol{x}_i$ の和として表す表し方が一意的なとき，和 $U_1 + \cdots + U_k$ を**直和**(direct sum)であるといい，

$$U_1 \oplus \cdots \oplus U_k \tag{3.2.10}$$

と記す．

注意 3.2.5 ベクトル空間を適当な部分空間の直和に分解して表すと見通しがよくなることが多い．その際，部分空間の和が直和になるための条件をいろいろな形で知っているとよい．よく知られているのは，部分空間 $U_1, \cdots, U_k$ の和 $U=U_1+\cdots+U_k$ に対して以下の 4 条件が等価であるということである．

(i) $\boldsymbol{x}_1+\cdots+\boldsymbol{x}_k=\boldsymbol{0}$，$\boldsymbol{x}_i \in U_i$ $(i=1,\cdots,k)$ となるのは $\boldsymbol{x}_i=\boldsymbol{0}$ $(i=1,\cdots,k)$ のときに限られる．

(ii) $U=U_1 \oplus \cdots \oplus U_k$.

(iii) $\dim U=\dim U_1+\cdots+\dim U_k$.

(iv) $U_1 \cap U_2=\{\boldsymbol{0}\}$，$(U_1+U_2) \cap U_3=\{\boldsymbol{0}\}, \cdots, (U_1+\cdots+U_{k-1}) \cap U_k=\{\boldsymbol{0}\}$.

略証 (i)→(ii)：$\boldsymbol{x}_1+\cdots+\boldsymbol{x}_k=\boldsymbol{y}_1+\cdots+\boldsymbol{y}_k$ で $\boldsymbol{x}_i \in U_i$，$\boldsymbol{y}_i \in U_i$ $(i=1,\cdots,k)$ ならば，$(\boldsymbol{x}_1-\boldsymbol{y}_1)+\cdots+(\boldsymbol{x}_k-\boldsymbol{y}_k)=\boldsymbol{0}$．$\boldsymbol{x}_i-\boldsymbol{y}_i \in U_i$ であるから $\boldsymbol{x}_i-\boldsymbol{y}_i=\boldsymbol{0}$，すなわち $\boldsymbol{x}_i=\boldsymbol{y}_i$ $(i=1,\cdots,k)$．(ii)→(iii)：$r_i=\dim U_i$ として U_i の基底 $B_i=\{\boldsymbol{e}_1^{(i)},\cdots,\boldsymbol{e}_{r_i}^{(i)}\}$ を選ぶと，$U=L(B_1 \cup \cdots \cup B_k)$ であるが，$\boldsymbol{x} \in U$ を $B_1 \cup \cdots \cup B_k$ の線形結合で表す表し方は(ii)により一意的であるから，$B_1 \cup \cdots \cup B_k$ は線形独立(たとえば $\sum_{i=1}^{k}\sum_{j=1}^{r_i}\lambda_j^{(i)}\boldsymbol{e}_j^{(i)}=\boldsymbol{0}$ はすべての $\lambda_j^{(i)}=0$ のときだけである)であり，したがって U の基底である．$r=\dim U=|B_1 \cup \cdots \cup B_k|=r_1+\cdots+r_k$．(iii)→(iv)：上と同様に B_i を定義すると，(iii)により $B_1 \cup \cdots \cup B_k$ が線形独立でなければならない．このことは，$B_1 \cup \cdots \cup B_i$ が線形独立で $U_1+\cdots+U_i$ の基底であることも意味する．$(B_1 \cup \cdots \cup B_i) \cup B_{i+1}$ が線形独立だから $(U_1+\cdots+U_i) \cap U_{i+1}=\{\boldsymbol{0}\}$ $(i=1,\cdots,k-1)$．(iv)→(i)：$\boldsymbol{x}_1+\cdots+\boldsymbol{x}_k=\boldsymbol{0}, \boldsymbol{x}_i \in U_i$ は $\boldsymbol{x}_k=-(\boldsymbol{x}_1+\cdots+\boldsymbol{x}_{k-1}) \in (U_1+\cdots+U_{k-1}) \cap U_k$ を意味するが，(iv)により $\boldsymbol{x}_k=\boldsymbol{0}$．したがって $\boldsymbol{x}_1+\cdots+\boldsymbol{x}_{k-1}=\boldsymbol{0}$．同様にして，$\boldsymbol{x}_i=\boldsymbol{0}$ $(i=k-1,\cdots,2,1)$． ■

3.2.3 補空間，商空間

ベクトル空間 V が二つの部分空間 U, W の直和として表されるとき，すなわち $V=U \oplus W$ であるとき，W は(V における)U の**補空間**(complementary subspace, complement)であるという(U は W の補空間である)．

$\dim V=n$ のとき，V の任意の部分空間 U に対して補空間 W が(一般には多数)存在し，$\dim W=n-\dim U$ である(注意 3.2.5 の(ii)，(iii))．

U が V の部分空間であるとする．$\boldsymbol{x}, \boldsymbol{y} \in V$ に対して

$$\boldsymbol{x}-\boldsymbol{y} \in U \iff \boldsymbol{x} \sim \boldsymbol{y} \tag{3.2.11}$$

によって V の中に定義される関係 $\sim$ は明らかに“同値関係”であり，関係 $\sim$ により V は同値類に分けられる．同値類の間に自然に“和”と“スカラー倍”とを定義すれば，同値類の集まりは一つのベクトル空間になることが容易に示さ

れる(代数学の言葉でいう“商系” $V/\sim$ である).

このベクトル空間のことを，V の U による**商空間**(quotient space)と呼び，V/U と記す．V における U の補空間の一つ W を選んで固定すると，各同値類は W の元 $\boldsymbol{y}$ を代表元として $\{\boldsymbol{y}+\boldsymbol{x} \mid \boldsymbol{x} \in U\}$ の形に表すことができる．したがって，

$$\dim(V/U)\,(=\dim W)=\dim V-\dim U \tag{3.2.12}$$

である．

3.3 線 形 写 像

3.3.1 線形写像と行列表現

U, V が K 上のベクトル空間であるとき，V から U への写像

$$A: V \longrightarrow U \tag{3.3.1}$$

が線形性の条件

(i) $A(\boldsymbol{x}+\boldsymbol{y})=A(\boldsymbol{x})+A(\boldsymbol{y}) \qquad (\boldsymbol{x}, \boldsymbol{y} \in V)$,　　(3.3.2)

(ii) $A(\lambda \boldsymbol{x})=\lambda A(\boldsymbol{x}) \qquad (\boldsymbol{x} \in V,\ \lambda \in K)$　　(3.3.3)

を満足するとき，A を V から U への**線形写像**(linear mapping, linear transformation)という(**一次写像**ともいう).

注意 3.3.1　(ii)において $\lambda=0$ とおけばわかるように，A は V の零ベクトルを U の零ベクトルに写す．また，$\boldsymbol{y}=A(\boldsymbol{x})$ なら $-\boldsymbol{y}=A(-\boldsymbol{x})$ である．代数学の言葉でいえば，線形写像は“準同型”である．

$\dim V_n=n$，$\dim U_m=m$ のとき，V_n, U_m にそれぞれ基底 $B=\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$，$\tilde{B}=\{\tilde{\boldsymbol{e}}_1, \cdots, \tilde{\boldsymbol{e}}_m\}$ を選んで固定する．$A(\boldsymbol{e}_j) \in U_m$ であるから

$$A(\boldsymbol{e}_j)=\sum_{i=1}^{m} a_{ij} \tilde{\boldsymbol{e}}_i \tag{3.3.4}$$

と書ける．すると，A の線形性から，V_n の任意のベクトル

$$\boldsymbol{x}=\sum_{j=1}^{n} x_j \boldsymbol{e}_j \tag{3.3.5}$$

の A による像は

$$A(\boldsymbol{x})=\sum_{i=1}^{m}\left(\sum_{j=1}^{n} a_{ij} x_j\right) \tilde{\boldsymbol{e}}_i \tag{3.3.6}$$

であるので，$A(\boldsymbol{x})$の座標表現

$$A(\boldsymbol{x})=\sum_{i=1}^{m}\tilde{x}_i\tilde{\boldsymbol{e}}_i \tag{3.3.7}$$

と比較すると

$$\tilde{x}_i=\sum_{j=1}^{n}a_{ij}x_j \tag{3.3.8}$$

であることがわかる．すなわち，選んだ基底 $B,\tilde{B}$ に関する座標を成分とする縦ベクトルで V_n,U_m のベクトルを表せば，線形写像 A は(3.3.4)の a_{ij} を成分とする $m\times n$ 型行列 $A=[a_{ij}]$によって

$$\begin{bmatrix}\tilde{x}_1\\ \vdots\\ \tilde{x}_m\end{bmatrix}=A\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix} \tag{3.3.9}$$

と表される．

以下では，線形写像 A をしばしばその行列表現と同一視して扱う．

l 次元，m 次元，n 次元のベクトル空間 U_l,V_m,W_n の間の二つの線形写像

$$B: W_n\longrightarrow V_m,\qquad A: V_m\longrightarrow U_l \tag{3.3.10}$$

の合成

$$C=A\cdot B: W_n\longrightarrow U_l \tag{3.3.11}$$

の行列表現は，$l\times m$ 行列 A と $m\times n$ 行列 B の積である $l\times n$ 行列 $C=A\cdot B$ であることも明らかであろう．

線形写像 $A: V_n\to U_m$ の基底 $B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\},\tilde{B}=\{\tilde{\boldsymbol{e}}_1,\cdots,\tilde{\boldsymbol{e}}_m\}$に関する表現が $m\times n$ 行列 $A=[a_{ij}]$であるとする．このとき，別の基底 $B'=\{\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n'\}$，$\tilde{B}'=\{\tilde{\boldsymbol{e}}_1',\cdots,\tilde{\boldsymbol{e}}_m'\}$に関する表現行列 A' の具体的な形を定めよう．基底変換の行列が，それぞれ，n 次と m 次の正則行列 $T=[t_{ik}]$，$S=[s_{jl}]$である：

$$[\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n']=[\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]T,\quad [\tilde{\boldsymbol{e}}_1',\cdots,\tilde{\boldsymbol{e}}_m']=[\tilde{\boldsymbol{e}}_1,\cdots,\tilde{\boldsymbol{e}}_m]S \tag{3.3.12}$$

とすると，$\boldsymbol{x}$ と $\tilde{\boldsymbol{x}}=A(\boldsymbol{x})$の座標成分表現は

$$\boldsymbol{x}=[\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}=[\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n']\begin{bmatrix}x_1'\\ \vdots\\ x_n'\end{bmatrix},$$
$$\tilde{\boldsymbol{x}}=[\tilde{\boldsymbol{e}}_1,\cdots,\tilde{\boldsymbol{e}}_m]\begin{bmatrix}\tilde{x}_1\\ \vdots\\ \tilde{x}_m\end{bmatrix}=[\tilde{\boldsymbol{e}}_1',\cdots,\tilde{\boldsymbol{e}}_m']\begin{bmatrix}\tilde{x}_1'\\ \vdots\\ \tilde{x}_m'\end{bmatrix} \tag{3.3.13}$$

であり，そして(3.3.9)，(3.1.13′)(の左側の式の S を T におきかえたもの)により

$$\tilde{\boldsymbol{x}}=[\tilde{\boldsymbol{e}}_1,\cdots,\tilde{\boldsymbol{e}}_m]\begin{bmatrix}\tilde{x}_1\\ \vdots\\ \tilde{x}_m\end{bmatrix}=[\tilde{\boldsymbol{e}}_1,\cdots,\tilde{\boldsymbol{e}}_m]A\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}$$
$$=[\tilde{\boldsymbol{e}}_1',\cdots,\tilde{\boldsymbol{e}}_m']S^{-1}AT\begin{bmatrix}x_1'\\ \vdots\\ x_n'\end{bmatrix}=[\tilde{\boldsymbol{e}}_1',\cdots,\tilde{\boldsymbol{e}}_m']\begin{bmatrix}\tilde{x}_1'\\ \vdots\\ \tilde{x}_m'\end{bmatrix}, \tag{3.3.14}$$

すなわち

$$\begin{bmatrix}\tilde{x}_1'\\ \vdots\\ \tilde{x}_m'\end{bmatrix}=S^{-1}AT\begin{bmatrix}x_1'\\ \vdots\\ x_n'\end{bmatrix} \tag{3.3.15}$$

を得る．このことは，基底変換(3.3.12)により，線形写像の表現行列は A から

$$A'=S^{-1}AT \tag{3.3.16}$$

に変換されることを意味している．(3.3.16)の形の変換を**同値変換**(equivalence transformation)と呼び，ある S, T を用いて同値変換により結ばれる二つの行列を**同値**(equivalent)であるということがある．

線形写像 A がベクトル空間 V_n からそれ自身への写像であるとき，表現行列は n 次正方行列であり，V_n の基底変換 S により表現行列は

$$A'=S^{-1}AS \tag{3.3.16′}$$

に変換される．このとき A(やA')の行集合と列集合とは同一である(厳密にいえば一対一対応がある(→注意 2.1.1))．(3.3.16′)の形の変換を**相似変換**(similarity transformation)と呼び，ある S を用いて相似変換により結ばれる二つの行列は**相似**(similar)であるといわれる．

3.3.2 線形写像のノルム

3.1.5 項で導入したベクトルのノルムに基づいて線形写像，あるいはその表現行列のノルムを定義することができる($K=\mathbb{R}$ または $\mathbb{C}$ とする)．ベクトルのノルムがそうであったように，線形写像，行列のノルムも，一般には，関係するベクトル空間の基底の選び方による．

ベクトル空間 U, V にそれぞれノルム $\|\cdot\|_U$, $\|\cdot\|_V$ が定められているとき，それらのノルムに対応する $A: V\to U$ の**作用素ノルム**(operator norm)(あるいは**自然ノルム**(natural norm))$\|A\|$は

$$\|A\|=\max_{\substack{\boldsymbol{x}\in V\\ \boldsymbol{x}\neq\boldsymbol{0}}}\frac{\|A\boldsymbol{x}\|_U}{\|\boldsymbol{x}\|_V}=\max_{\substack{\boldsymbol{x}\in V\\ \|\boldsymbol{x}\|_V=1}}\|A\boldsymbol{x}\|_U \tag{3.3.17}$$

で定義される．この $\|A\|$ は，$m\times n$ 行列の全体 $\mathcal{M}_{m,n}$ を(行列の加法とスカラー倍についての)$m\cdot n$ 次元のベクトル空間とみなしたとき(→演習問題 3.1(v))，ベクトルのノルムの満たすべき 3 条件(3.1.5 項の(1)，(2)，(3))を満たしている．

(1′) $\|A\|\geqq 0$； $\|A\|=0 \iff A=O_{m,n}$,

(2′) $\|\lambda A\|=|\lambda|\cdot\|A\| \qquad (\lambda\in K)$,

(3′) A, B が同じ型のとき $\|A+B\|\leqq\|A\|+\|B\|$.

略証 (1′)，(2′)は明らか．(3′)については

$$\begin{aligned}\|A+B\|&=\max_{\substack{\boldsymbol{x}\in V\\ \|\boldsymbol{x}\|_V=1}}\|(A+B)\boldsymbol{x}\|_U=\max_{\substack{\boldsymbol{x}\in V\\ \|\boldsymbol{x}\|_V=1}}\|A\boldsymbol{x}+B\boldsymbol{x}\|_U\\ &\leqq\max_{\substack{\boldsymbol{x}\in V\\ \|\boldsymbol{x}\|_V=1}}(\|A\boldsymbol{x}\|_U+\|B\boldsymbol{x}\|_U)\leqq\max_{\substack{\boldsymbol{x}\in V\\ \|\boldsymbol{x}\|_V=1}}\|A\boldsymbol{x}\|_U+\max_{\substack{\boldsymbol{x}\in V\\ \|\boldsymbol{x}\|_V=1}}\|B\boldsymbol{x}\|_U\\ &=\|A\|+\|B\|.\end{aligned}$$

■

ベクトル空間 U_l, V_m, W_n に，それぞれ，ノルム $\|\cdot\|_U, \|\cdot\|_V, \|\cdot\|_W$ が定められていて，$B: W\to V$，$A: V\to U$ であるとき，$A\cdot B: W\to U$ のノルムは，

(4′) $\|A\cdot B\|\leqq\|A\|\cdot\|B\|$

を満たす．

略証 $B\neq O_{m,n}$ として，

$$\|A\cdot B\|=\max_{\substack{x\in W\\ x\neq 0}}\frac{\|A\cdot Bx\|_U}{\|x\|_W}=\max_{\substack{x\in W\\ x\neq 0}}\left(\frac{\|A(Bx)\|_U}{\|Bx\|_V}\cdot\frac{\|Bx\|_V}{\|x\|_W}\right)$$

$$\leqq\left(\max_{\substack{x\in W\\ x\neq 0}}\frac{\|A(Bx)\|_U}{\|Bx\|_V}\right)\cdot\left(\max_{\substack{x\in W\\ x\neq 0}}\frac{\|Bx\|_V}{\|x\|_W}\right)\leqq\left(\max_{\substack{y\in V\\ y\neq 0}}\frac{\|Ay\|_U}{\|y\|_V}\right)\cdot\left(\max_{\substack{x\in W\\ x\neq 0}}\frac{\|Bx\|_V}{\|x\|_W}\right)$$

$$=\|A\|\cdot\|B\|.$$ ■

一般に行列に非負実数値を対応させる関数で，(1′)，(2′)，(3′)，(4′)を満たすものを**行列ノルム**(matrix norm)と呼ぶ．

例 3.3.1　$A: V_n\to U_m$ に対して，$\|\cdot\|_U$ も $\|\cdot\|_V$ も絶対値和ノルムあるいは最大値ノルム(例 3.1.2)のとき，$\|A\|$ を $\|A\|_1$ あるいは $\|A\|_\infty$ と書くが，それらは具体的に

$$\|A\|_1=\max_{j=1,\cdots,n}\sum_{i=1}^{m}|a_{ij}|, \tag{3.3.18a}$$

$$\|A\|_\infty=\max_{i=1,\cdots,m}\sum_{j=1}^{n}|a_{ij}| \tag{3.3.18b}$$

と表される．実際，$\|Ax\|_1=\sum_{i=1}^{m}\left|\sum_{j=1}^{n}a_{ij}x_j\right|\leqq\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|\cdot|x_j|=\sum_{j=1}^{n}|x_j|\sum_{i=1}^{m}|a_{ij}|\leqq\left(\sum_{j=1}^{n}|x_j|\right)\cdot\left(\max_{j=1,\cdots,n}\sum_{i=1}^{m}|a_{ij}|\right)=\|x\|_1\cdot\left(\max_{j=1,\cdots,n}\sum_{i=1}^{m}|a_{ij}|\right)$．したがって，$\|A\|_1\leqq\max_j\sum_i|a_{ij}|$．$\max_j\sum_i|a_{ij}|=\sum_i|a_{ij_0}|$ のとき，$x_j=0$ $(j\neq j_0)$，$x_{j_0}=1$ とおけば，上の不等号はみな等号で成立するから，$\|Ax\|_1=\sum_i|a_{ij_0}|$ であり，(3.3.18a)が得られる．

同様に，$\|Ax\|_\infty=\max_{i=1,\cdots,m}\left|\sum_{j=1}^{n}a_{ij}x_j\right|\leqq\max_i\sum_j|a_{ij}|\cdot|x_j|\leqq\max_i\left(\sum_j|a_{ij}|\right)\cdot\max_j|x_j|=\left(\max_i\sum_j|a_{ij}|\right)\cdot\|x\|_\infty$．したがって $\|A\|_\infty\leqq\max_i\sum_j|a_{ij}|$．$\max_i\sum_j|a_{ij}|=\sum_j|a_{i_0j}|$ のとき，$x_j=0$ $(a_{i_0j}=0$ のとき$)$，$x_j=\bar{a}_{i_0j}/|a_{i_0j}|$ $(a_{i_0j}\neq 0$ のとき$)$とおけば，上の不等号はみな等号となり，$\|x\|_\infty=1$ であるから，$\|Ax\|_\infty=\sum_j|a_{i_0j}|$ であり，(3.3.18b)が得られる． □

例 3.3.2　$\|\cdot\|_U$ も $\|\cdot\|_V$ も 2 乗ノルム(3.1.5 項)のときには$(\|Ax\|_2)^2=(Ax)^{\mathrm{H}}(Ax)=x^{\mathrm{H}}A^{\mathrm{H}}Ax=x^{\mathrm{H}}(A^{\mathrm{H}}A)x$ であるが，このことは，後に 6.9 節において述べる特異値の概念と結びつけると，

$$\|A\|_2=A\text{ の最大特異値} \tag{3.3.19}$$

といい表せる． □

例 3.3.3　作用素ノルム以外に行列ノルムでよく利用されるものに **Frobenius**

(フロベニウス)**ノルム**

$$\|A\|_F=\left(\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2\right)^{1/2} \tag{3.3.20}$$

がある．これは $m\times n$ 行列の全体 $\mathcal{M}_{m,n}$ をベクトル空間とみなしたときの 2 乗ノルムであるから，行列ノルムの条件(1′)，(2′)，(3′)を満たす．(4′)を満たすことは次のようにして示される．

$$\|AB\|_F{}^2=\sum_i\sum_j\left|\sum_k a_{ik}b_{kj}\right|^2\leqq\sum_i\sum_j\left(\sum_k|a_{ik}||b_{kj}|\right)^2$$

であるが，Cauchy-Schwarz の不等式(2.16.13)により

$$\leqq\sum_i\sum_j\left(\sum_k|a_{ik}|^2\right)\left(\sum_l|b_{lj}|^2\right)=\left(\sum_i\sum_k|a_{ik}|^2\right)\left(\sum_j\sum_l|b_{lj}|^2\right)$$
$$=\|A\|_F{}^2\cdot\|B\|_F{}^2.$$

例 3.3.2，(3.3.19)と関連づけていえば，

$$\|A\|_F{}^2=A\text{ の特異値の 2 乗の和} \tag{3.3.21}$$

である(6.9 節で述べる)．したがって

$$\|A\|_2\leqq\|A\|_F \tag{3.3.22}$$

である． □

例 3.3.4 ユニタリ行列 Q については $\|Q\boldsymbol{x}\|_2=\|\boldsymbol{x}\|_2$ であるから(→(2.14.26))，$Q:V\to V$ の 2 乗ノルムは $\|Q\|_2=1$ である． □

注意 3.3.2 写像(あるいはその表現行列)のあるノルム $\|\cdot\|$ に対して，

$$\mathrm{cond}\,A=\|A\|\cdot\|A^{-1}\| \tag{3.3.23}$$

で定義される量を A の**条件数**(condition number)と呼ぶ．ノルムの種類を明示するために，cond_∞，$\mathrm{cond}_p\,(p\geqq1)$，$\mathrm{cond}_F$ などと記すこともある．条件数は，線形方程式系の数値解法や固有値問題の数値解法において重要な役割を果たす(条件数が大きいほど数値的には扱い難いとされている)．

対角行列 $A=\mathrm{diag}[a_{11},\cdots,a_{nn}]$ の条件数は $\mathrm{cond}_p\,A=\max_i|a_{ii}|/\min_i|a_{ii}|$ である．ユニタリ行列 Q の条件数は $\mathrm{cond}_2\,Q=1$ である．

3.3.3 零空間と像空間

線形写像 $A:V_n\to U_m$ に対して，$A\boldsymbol{x}=\boldsymbol{0}$ となるようなベクトル $\boldsymbol{x}(\in V_n)$の全体は，明らかに，V_n の部分空間をなす．これを A の**零空間**(null space)，ある

いは**核空間**(kernel (space))といい，Ker A と記す．すなわち

$$\mathrm{Ker}\,A=\{\boldsymbol{x}\in V_n \mid A\boldsymbol{x}=\boldsymbol{0}\}. \tag{3.3.24}$$

ある $\boldsymbol{x}\in V_n$ の A による像となるような U_m のベクトル $\boldsymbol{y}=A\boldsymbol{x}$ の全体も，明らかに，U_m の部分空間をなす．これを A の**像空間**(image(space))といい Im A と記す．すなわち

$$\mathrm{Im}\,A=\{\boldsymbol{y}\in U_m \mid \exists\,\boldsymbol{x}\in V_n: \boldsymbol{y}=A\boldsymbol{x}\}. \tag{3.3.25}$$

注意 3.3.3　A を行列表現したときの列ベクトルを $\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot n}$ と記すと，(3.3.25)は，Im A がある $\boldsymbol{x}=[x_1,\cdots,x_n]^{\mathrm{T}}$ を用いて

$$\boldsymbol{y}=\sum_{j=1}^{n}x_j\boldsymbol{a}_{\cdot j} \tag{3.3.26}$$

と表される U_m のベクトルであるということを述べている．すなわち，3.2.1 項の言葉を用いると，Im A は A の列ベクトルにより生成される U_m の部分空間である：

$$\mathrm{Im}\,A=L(\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot n}). \tag{3.3.27}$$

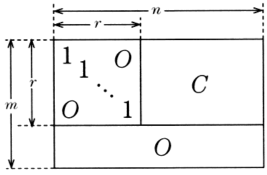
したがって，3.1.4 項により，

$$\dim(\mathrm{Im}\,A)=\mathrm{rank}\,A \tag{3.3.28}$$

である(左辺の A は“写像”，右辺の A は“行列”とみなして)．特に，rank $A=m$ なら Im $A=U_m$．

このように，行列 A の階数は A の列ベクトルの張るベクトル空間の次元に等しい．rank A^{T}=rank A であるから(2.13.1 項)，それは A の行ベクトルの張る空間の次元にも等しい．rank A を写像 A の階数(rank)ということもある．

注意 3.3.4　U_m の基底を適当に選び V_n の基底のベクトルの順序を適当に選ぶことによって写像 A の表現行列を(3.3.29)の形にすることができる．すなわち，適当な m 次正則行列 S と n 次置換行列 P を用いて SAP を(3.3.29)の形にすることができる(→2.15.1 項，基本変形(2.15.1))．



(3.3.29)

ここで rank $A=r$ である．そこで，$\boldsymbol{x}=[x_1,\cdots,x_n]^{\mathrm{T}}$ が $A\boldsymbol{x}=\boldsymbol{0}$ を満たすための必要十分条件は

$$x_i+\sum_{j=1}^{n-r}c_{ij}x_{r+j}=0 \qquad (i=1,\cdots,r) \tag{3.3.30}$$

(部分行列 C の要素を $c_{ij}(i=1,\cdots,r\,;\,j=1,\cdots,n-r)$ と記した)．このときの V_n の基底を $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ とすれば，(3.3.30)は

$$\boldsymbol{x}=\sum_{i=1}^{r}\left(-\sum_{j=1}^{n-r}c_{ij}x_{r+j}\right)\boldsymbol{e}_i+\sum_{j=1}^{n-r}x_{r+j}\boldsymbol{e}_{r+j}$$

$$=\sum_{j=1}^{n-r}x_{r+j}\left(\boldsymbol{e}_{r+j}-\sum_{i=1}^{r}c_{ij}\boldsymbol{e}_i\right) \tag{3.3.31}$$

とも表される．ここで，$x_{r+j}(j=1,\cdots,n-r)$ は任意であるから，(3.3.31)は

$$\left\{\boldsymbol{e}_{r+1}-\sum_{i=1}^{r}c_{i1}\boldsymbol{e}_i,\cdots,\boldsymbol{e}_n-\sum_{i=1}^{r}c_{i,n-r}\boldsymbol{e}_i\right\} \tag{3.3.32}$$

が Ker A の一つの基底であること，したがって，

$$\dim(\mathrm{Ker}\,A)=n-\mathrm{rank}\,A \tag{3.3.33}$$

すなわち，

$$\dim(\mathrm{Ker}\,A)+\dim(\mathrm{Im}\,A)=n \tag{3.3.34}$$

であることも意味している．特に，rank $A=n$ なら Ker $A=\{\mathbf{0}\}$．

写像 A の零空間の次元(3.3.33)を写像 A の**零度**(nullity)と呼び null A と記すことがある．

3.3.4 不変部分空間

ベクトル空間 V からそれ自身への線形写像(あるいはその行列表現) A：$V\to V$ と V の部分空間 U が

$$\text{すべての } \boldsymbol{x}\in U \text{ に対して } A\boldsymbol{x}\in U \tag{3.3.35}$$

を満たすとき，U は A の**不変部分空間**(invariant subspace)であるという．このとき，A は U における線形写像 A：$U\to U$ であるともみなせる(これを A の U への**制限**(restriction)と呼び，$A|_U$ と記したりすることもある)．V 自身および $\{\mathbf{0}\}$ は自明な不変部分空間である．

注意 3.3.5　U が A の不変部分空間ならば，U は任意の x の多項式 $f(x)$ から導かれる行列多項式(→注意 2.2.3) $f(A)$ の不変部分空間でもある．

U_1 と U_2 が共に A の不変部分空間であるならば，それらの共通部分 $U_1\cap U_2$ も和 U_1+U_2 も A の不変部分空間である．

U が A の不変部分空間であるとき，U の補空間(→3.2.3 項)の一つを W として，U と W に基底 $B_0=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_m\}, B_1=\{\boldsymbol{e}_{m+1},\cdots,\boldsymbol{e}_n\}$ を選べば($\dim U=m$, $\dim W=n-m$)，$B_0\cup B_1$ は V の基底であるが(→3.2.2，3 項)，この基底に関する A の表現行列が

$$S^{-1}AS=\left[\begin{array}{c|c} \overset{\leftarrow m \rightarrow}{A_{00}} & \overset{\leftarrow n-m \rightarrow}{A_{01}} \\ \hline O & A_{11} \end{array}\right]\begin{array}{l} \updownarrow m \\ \updownarrow n-m \end{array} \tag{3.3.36}$$

という形になることは明らかであろう．（S はもとの基底と $B_0 \cup B_1$ とを結ぶ基底変換の行列．）ここで A_{00} は $A|_U$ の一つの表現行列である．

この形に表現しておけば，$\boldsymbol{x} \in V$ は $\boldsymbol{x}=[\boldsymbol{x}_0{}^\top, \boldsymbol{x}_1{}^\top]^\top$ という形の座標表現をもつが，$\boldsymbol{x} \in U$ なら $\boldsymbol{x}_1=\boldsymbol{0}$，$\boldsymbol{x} \in W$ なら $\boldsymbol{x}_0=\boldsymbol{0}$ である．そこで，$\boldsymbol{y}=A\boldsymbol{x}$ は

$$\boldsymbol{y}_0=A_{00}\boldsymbol{x}_0+A_{01}\boldsymbol{x}_1, \qquad \boldsymbol{y}_1=A_{11}\boldsymbol{x}_1 \tag{3.3.37}$$

と表される．$\boldsymbol{x} \in U$ なら $\boldsymbol{x}_1=\boldsymbol{0}$ で $\boldsymbol{y}_0=A_{00}\boldsymbol{x}_0$ となるが，この A_{00} が U における A の表現になっているのは上に述べた通りである．ところで，$\boldsymbol{x} \in W$ なら $\boldsymbol{x}_0=\boldsymbol{0}$ であるから，$\boldsymbol{y}_0=A_{01}\boldsymbol{x}_1$，$\boldsymbol{y}_1=A_{11}\boldsymbol{x}_1$ となる．W も A の不変部分空間であるためには $A_{01}=O_{m,n-m}$ でなければならない．しかるに，この表現からも明らかなように，任意の $\boldsymbol{x} \in V$ に対して $\boldsymbol{y}_1=A_{11}\boldsymbol{x}_1$ であるので，A_{11} は商空間 V/U の上に A が定義する線形写像の表現行列であるとみなせる．これは補空間 W の選び方によらない（$\boldsymbol{x}'-\boldsymbol{x} \in U, \boldsymbol{y}'=A\boldsymbol{x}', \boldsymbol{y}=A\boldsymbol{x}$ なら $\boldsymbol{y}'-\boldsymbol{y} \in U$ であるから）．この写像は，A の V/U の上への**縮約**（contraction）と呼ばれることがある．

U も W も A の不変部分空間であるときには $A_{01}=O$ であるが，このことを簡単のため

$$A=A_{00}\dot{+}A_{11} \tag{3.3.38}$$

と書き表すことにする．

注意 3.3.6 $A: V \to V$ に反傾な変換 $A^\top: V^{\mathrm{d}} \to V^{\mathrm{d}}$（次節 3.4.3 項を見よ）を考えると，$A_{11}{}^\top$ が V^{d} の部分空間への制限，$A_{00}{}^\top$ が商空間への縮約となる．

3.4 双対ベクトル空間

3.4.1 線形形式，双対空間，スカラー積

V を K 上のベクトル空間とする．K 自身も K 上の 1 次元ベクトル空間であるから，線形写像 $f: V \to K$ ［これを**線形形式**（あるいは**一次形式**；linear form）という］の全体（これを V^{d} と記す）を考えることができる．V^{d} は写像の和とスカラー倍の演算に関して一つのベクトル空間をなす．これを V の**双対（ベクトル）空間**（dual (vector) space）と呼ぶ．（この具体的イメージは次の 3.4.2

項で与える.)

$\boldsymbol{x}(\in V)$の$f(\in V^{\mathrm{d}})$による像は$f(\boldsymbol{x})(\in K)$であるが，この$f(\boldsymbol{x})$を$\boldsymbol{x}\in V$と$\boldsymbol{f}\in V^{\mathrm{d}}$の一種の"積"とみなして$\langle\boldsymbol{x},\boldsymbol{f}\rangle$のように表し，$\boldsymbol{x}$と$\boldsymbol{f}$($f$がベクトル空間の元，すなわちベクトルであるので太文字を使う)との**スカラー積**(scalar product)と呼んだりもする.

3.4.2 双対基底

Vの一つの基底を$B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$とする. Kをベクトル空間とみなし，その基底として$\{1\}$を選ぶと，線形形式$f\colon V\to K$の表現行列は$1\times n$型で$\boldsymbol{f}=[z_1,\cdots,z_n]\ (z_j\in K)$のように表せる. このとき

$$f(\boldsymbol{x})=\langle\boldsymbol{x},\boldsymbol{f}\rangle=\sum_{j=1}^{n}z_jx_j \tag{3.4.1}$$

となる. $\boldsymbol{f}$を横ベクトル，$\boldsymbol{x}=[x_1,\cdots,x_n]^{\mathrm{T}}$を縦ベクトルとみなすと，(3.4.1)は行列算における積(2.2.2項)$\boldsymbol{f}\cdot\boldsymbol{x}$とも同一視できる.

いま，特に，$\boldsymbol{e}_i^{\mathrm{d}}(\in V^{\mathrm{d}})$として

$$\boldsymbol{e}_i^{\mathrm{d}}(\boldsymbol{e}_j)=\langle\boldsymbol{e}_j,\boldsymbol{e}_i^{\mathrm{d}}\rangle=\delta_{ij}=\begin{cases}1 & (i=j\text{ のとき})\\ 0 & (i\neq j\text{ のとき}),\end{cases} \tag{3.4.2}$$

すなわち

$$\boldsymbol{e}_i^{\mathrm{d}}(\boldsymbol{x})=x_i \tag{3.4.2$'$}$$

であるような線形形式を選ぶと，任意の線形形式は，(3.4.1)により

$$\boldsymbol{f}=z_1\boldsymbol{e}_1^{\mathrm{d}}+\cdots+z_n\boldsymbol{e}_n^{\mathrm{d}} \tag{3.4.3}$$

と一意的に表せる. このことは$B^{\mathrm{d}}=\{\boldsymbol{e}_1^{\mathrm{d}},\cdots,\boldsymbol{e}_n^{\mathrm{d}}\}$が$V^{\mathrm{d}}$の一つの基底であることを示している. B^{d}のことを，Vの基底Bに対応する**双対基底**(dual basis)と呼ぶ.

注意 3.4.1 以上のことから自然に

$$\dim V=\dim V^{\mathrm{d}} \tag{3.4.4}$$

であることも示されたことになる. また，基底と双対基底の関係が対称であるから，$(V^{\mathrm{d}})^{\mathrm{d}}=V$となることも明らかであろう.

注意 3.4.2 Vの基底を$B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$から$B'=\{\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n'\}$に変換するとき，

$$[\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n']=[\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]S \tag{3.4.5}$$

に対応する双対基底は$B^{\mathrm{d}}=\{\boldsymbol{e}_1^{\mathrm{d}},\cdots,\boldsymbol{e}_n^{\mathrm{d}}\}$から$B'^{\mathrm{d}}=\{\boldsymbol{e}_1'^{\mathrm{d}},\cdots,\boldsymbol{e}_n'^{\mathrm{d}}\}$に変換されるが，$B'$

と B'^{d} も(3.4.2)を満たさなければならないから，B^{d} と B'^{d} とは

$$\begin{bmatrix} \boldsymbol{e}_1^{\mathrm{d}} \\ \vdots \\ \boldsymbol{e}_n^{\mathrm{d}} \end{bmatrix} = S \begin{bmatrix} \boldsymbol{e}_1'^{\mathrm{d}} \\ \vdots \\ \boldsymbol{e}_n'^{\mathrm{d}} \end{bmatrix} \tag{3.4.6}$$

という形の変換を受ける．[(3.4.5)と(3.4.6)も互いに**反傾な関係**にある．]

また，V のベクトル $\boldsymbol{x}$ の基底 B に関する座標成分は，

$$\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{e}_j \tag{3.4.7}$$

の両辺の双対基底 B^{d} のベクトル $\boldsymbol{e}_i^{\mathrm{d}}$ とのスカラー積を取ることによって，

$$x_i = \langle \boldsymbol{x}, \boldsymbol{e}_i^{\mathrm{d}} \rangle \tag{3.4.8}$$

のように表すことができる(基底変換と座標成分の変換が反傾な関係にあることと平仄が合っている)．

注意 3.4.3　V のベクトルと V^{d} のベクトルとの“スカラー積”と，V の二つのベクトルの“内積”(3.5.4 項)とを混同してはならない．

3.4.3　反傾変換，双対変換，対称変換

ベクトル空間 V_n から U_m の双対空間 U_m^{d} への線形写像 $A: V_n \to U_m^{\mathrm{d}}$ を考える．V_n と U_m に基底を定めれば，U_m^{d} にも双対基底が定まり，A は $m \times n$ 行列 $A = [a_{ij}]$ で表現される．$\boldsymbol{y} = A\boldsymbol{x}$ ($\boldsymbol{y} \in U_m^{\mathrm{d}}$, $\boldsymbol{x} \in V_n$)の成分の間の関係式は

$$y_i = \sum_{j=1}^{n} a_{ij} x_j \qquad (i = 1, \cdots, m) \tag{3.4.9}$$

である．任意の $\boldsymbol{z} \in U_m$ に対して

$$\langle \boldsymbol{z}, A\boldsymbol{x} \rangle = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} z_i x_j \tag{3.4.10}$$

であるが，(3.4.10)は，$\boldsymbol{z}$ を固定したとき，$\boldsymbol{x}(\in V_n)$ に対して $\langle \boldsymbol{z}, A\boldsymbol{x} \rangle (\in K)$ を対応させる線形形式 $\boldsymbol{f}(\in V_n^{\mathrm{d}})$ であるともみなせる．このようにして，$\boldsymbol{z}(\in U_m)$ に $\boldsymbol{f}(\in V_n^{\mathrm{d}})$ を対応づける一つの写像(容易に確かめられるようにこれは線形写像である)が得られる．この写像の同じ基底による表現行列が A の転置 A^{T} であることは明らかである．そこでこの写像自身も A^{T} と記すことにする．

$$A: V_n \to U_m^{\mathrm{d}}, \qquad A^{\mathrm{T}}: U_m \to V_n^{\mathrm{d}}. \tag{3.4.11}$$

A^{T} は A に**反傾**(contragredient)であるという．

$\boldsymbol{w} = A^{\mathrm{T}}\boldsymbol{z}$ の成分による表現は

$$w_j = \sum_{i=1}^{m} a_{ij} z_i \tag{3.4.12}$$

であり，$\boldsymbol{x}(\in V_n)$とのスカラー積を取ると

$$\langle \boldsymbol{x}, A^{\top}\boldsymbol{z}\rangle = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}z_i x_j = \langle \boldsymbol{z}, A\boldsymbol{x}\rangle \qquad (3.4.13)$$

となる[(3.4.13)の最左辺は V_n と V_n^{d} の間のスカラー積，最右辺は U_m と U_m^{d} の間のスカラー積であることに注意].

特に $V=U$ であるとき，$A: V_n \to V_n^{\mathrm{d}}$ に対して $A^{\top}: V_n \to V_n^{\mathrm{d}}$ は A に**双対**(dual)であるという．両者が一致するとき，$A=A^{\top}$ は**対称**(symmetric)であるという(対称変換の表現行列は，V_n の基底の双対基底を V_n^{d} の基底として選べば，対称行列である).

3.5　双線形形式と二次形式

3.5.1　双線形形式

3.4.3項において見たように，V_n から U_m^{d}(U_m の双対空間)への線形写像 $A: V_n \to U_m^{\mathrm{d}}$ に対して

$$g: (\boldsymbol{x}, \boldsymbol{z}) \mapsto \langle \boldsymbol{z}, A\boldsymbol{x}\rangle \qquad (3.5.1)$$

(左辺の$(\boldsymbol{x}, \boldsymbol{z})$は $\boldsymbol{x}$ と $\boldsymbol{z}$ の対$(\in V_n \times U_m)$を表す)によって定義される写像 $g: V_n \times U_m \to K$ は，

(i)　$\boldsymbol{z} \in U_m$ を固定すれば $g(\cdot, \boldsymbol{z}): V_n \to K$ が線形形式，

(ii)　$\boldsymbol{x} \in V_n$ を固定すれば $g(\boldsymbol{x}, \cdot): U_m \to K$ が線形形式

となっている．

一般に，(i)，(ii)を満たす $g: V_n \times U_m \to K$ のことを**双線形形式**(**双一次形式**；bilinear form)と呼ぶ．V_n, U_m にそれぞれ基底$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}, \{\tilde{\boldsymbol{e}}_1, \cdots, \tilde{\boldsymbol{e}}_m\}$を選んで固定し

$$a_{ij} = g(\boldsymbol{e}_j, \tilde{\boldsymbol{e}}_i) \qquad (i=1, \cdots, m; j=1, \cdots, n) \qquad (3.5.2)$$

により $m \times n$ 行列 $A=[a_{ij}]$を定め，$A: V_n \to U_m^{\mathrm{d}}$ を考えれば，$\boldsymbol{x} \in V_n$，$\boldsymbol{z} \in U_m$ に対して

$$g(\boldsymbol{x}, \boldsymbol{z}) = \langle \boldsymbol{z}, A\boldsymbol{x}\rangle \qquad (3.5.3)$$

と表せることは明らかである($\langle \cdot, \cdot \rangle$は U_m と U_m^{d} の間のスカラー積)．行列 A は双線形形式のそれぞれの基底に関する“係数行列”である．

注意 3.5.1 (3.5.3)は A に反傾な変換 $A^{\mathrm{T}}: U_m \to V_n^{\mathrm{d}}$ を用いて $g(\boldsymbol{x},\boldsymbol{z})=\langle \boldsymbol{x}, A^{\mathrm{T}}\boldsymbol{z}\rangle$ ($\langle \cdot , \cdot \rangle$は V_n と V_n^{d} の間のスカラー積)とも表せる(→(3.4.13)).

V_n の基底を$\{\boldsymbol{e}_1', \cdots, \boldsymbol{e}_n'\}$に，$U_m$ の基底を$\{\tilde{\boldsymbol{e}}_1', \cdots, \tilde{\boldsymbol{e}}_m'\}$に変換する：

$$[\boldsymbol{e}_1', \cdots, \boldsymbol{e}_n'] = [\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n]\,T, \quad [\tilde{\boldsymbol{e}}_1', \cdots, \tilde{\boldsymbol{e}}_m'] = [\tilde{\boldsymbol{e}}_1, \cdots, \tilde{\boldsymbol{e}}_m]\,S \tag{3.5.4}$$

(T は n 次正則，S は m 次正則)．すると，

$$a_{ij}' = g(\boldsymbol{e}_j', \tilde{\boldsymbol{e}}_i') = g\left(\sum_{l=1}^{n} t_{lj}\boldsymbol{e}_l, \sum_{k=1}^{m} s_{ki}\tilde{\boldsymbol{e}}_k\right)$$

$$= \sum_{k=1}^{m}\sum_{l=1}^{n} s_{ki}a_{kl}t_{lj} \tag{3.5.5}$$

であるから，双線形形式の係数行列は A から

$$A' = S^{\mathrm{T}}AT \tag{3.5.6}$$

に変換される．

特に，$V_n = U_m$ $(n=m)$であって，基底も V_n と U_m とで同一のものを用いるときには，A は正方行列で，$S=T$ であり，基底変換に伴う双線形形式の係数行列の変換は

$$A' = S^{\mathrm{T}}AS \tag{3.5.7}$$

の形を取る．(3.5.7)の形の変換を**合同変換**(congruence transformation)と呼ぶ．

注意 3.5.2 V の基底変換に伴って，線形写像 $A: V \to V$ の表現行列が受ける変換 $A' = S^{-1}AS$(相似変換；3.3.1 項，(3.3.16′))と双線形形式$\langle \boldsymbol{z}, A\boldsymbol{x}\rangle$($A: V \to V^{\mathrm{d}}$)の係数行列が受ける変換 $A' = S^{\mathrm{T}}AS$(合同変換；(3.5.7))との相異に注意．S が“直交行列”であるときに限って，両者が一致する．

3.5.2 複素共役線形形式，複素双線形形式

$K=\mathbb{C}$，すなわち複素数体の上のベクトル空間については，複素共役量を考える必要がある場合が少なくない．そのような場合，3.4 節における線形形式，双対空間，スカラー積，反傾変換，などなどの概念，および本節で扱う双線形形式，二次形式，内積，などの概念を以下のように修正したものがよく使われる．

$f: V \to K$ が

$$\left.\begin{aligned} f(\boldsymbol{x}+\boldsymbol{y}) &= f(\boldsymbol{x})+f(\boldsymbol{y}) \qquad (\boldsymbol{x},\boldsymbol{y}\in V), \\ f(\lambda\boldsymbol{x}) &= \bar{\lambda}f(\boldsymbol{x}) \qquad (\boldsymbol{x}\in V, \lambda\in K) \end{aligned}\right\} \tag{3.5.8}$$

を満たすとき，f を**複素共役線形形式**と呼ぶことにする((3.5.8)の第2式の右辺に，λ でなくその共役複素数 $\bar{\lambda}$ が現れることに注意)．

複素共役線形形式の全体がベクトル空間をなすことは容易に示せる（このときのスカラー倍 λf には複素共役は使わない)．この空間を V^* と記し，**共役双対空間**と呼ぶ．V に基底 $B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$を選び，特別な複素共役線形形式

$$\boldsymbol{e}_i^*(\boldsymbol{e}_j)=\delta_{ij} \qquad (i,j=1,\cdots,n) \tag{3.5.9}$$

を定める．$\boldsymbol{f}\in V^*$，$\boldsymbol{x}=\sum_{j=1}^{n}x_j\boldsymbol{e}_j$ に対して，

$$\boldsymbol{f}(\boldsymbol{e}_j)=z_j \tag{3.5.10}$$

であれば

$$\boldsymbol{f}(\boldsymbol{x})=\sum_{j=1}^{n}z_j\bar{x}_j \tag{3.5.11}$$

となる．(3.5.11)を $\boldsymbol{x}\in V$ と $\boldsymbol{f}\in V^*$ の**複素スカラー積**と呼び，前節と同一の記号$\langle\boldsymbol{x},\boldsymbol{f}\rangle$で表す．このとき

$$\boldsymbol{f}=\sum_{j=1}^{n}z_j\boldsymbol{e}_j^* \tag{3.5.12}$$

であって，$B^*=\{\boldsymbol{e}_1^*,\cdots,\boldsymbol{e}_n^*\}$が V^* の基底をなし，B^* を B の**共役双対基底**と呼ぶ．

注意 3.5.3 複素スカラー積については，$\boldsymbol{x}\in V$ に対して$\langle\boldsymbol{x},\cdot\rangle$：$V^*\to K$ が(普通の)線形形式，$\boldsymbol{f}\in V^*$ に対して$\langle\cdot,\boldsymbol{f}\rangle$：$V\to K$ が複素共役線形形式となる．

線形写像 A：$V\to U^*$ に関連して定義される

$$g(\boldsymbol{x},\boldsymbol{z})=\langle\boldsymbol{z},A\boldsymbol{x}\rangle \qquad (\boldsymbol{z}\in U,\boldsymbol{x}\in V) \tag{3.5.13}$$

は，$g(\boldsymbol{x},\cdot)$：$U\to K$ が線形形式，$g(\cdot,\boldsymbol{z})$：$V\to K$ が複素共役線形形式である．このような写像 g：$V\times U\to K$ のことを**複素双線形形式**と呼ぶことにする．(3.5.13)は，成分で表せば，$m\times n$ 行列 $A=[a_{ij}]$を係数行列として，

$$\langle\boldsymbol{z},A\boldsymbol{x}\rangle=\sum_{i=1}^{m}\sum_{j=1}^{n}a_{ij}\bar{z}_ix_j. \tag{3.5.14}$$

$\boldsymbol{x},\boldsymbol{z}$ を，それぞれ，縦ベクトル$[x_1,\cdots,x_n]^{\top}$, $[z_1,\cdots,z_m]^{\top}$ とすれば

$$\langle\boldsymbol{z},A\boldsymbol{x}\rangle=\boldsymbol{z}^{\mathrm{H}}A\boldsymbol{x} \tag{3.5.14'}$$

である(この場合 $\boldsymbol{z}^{\mathrm{H}}=\overline{(\boldsymbol{z}^{\mathrm{T}})}$).

(3.5.14)の複素共役を取ると

$$g^*(\boldsymbol{z},\boldsymbol{x})\equiv\overline{\langle\boldsymbol{z},A\boldsymbol{x}\rangle}=\boldsymbol{x}^{\mathrm{H}}A^{\mathrm{H}}\boldsymbol{z}=\langle\boldsymbol{x},A^{\mathrm{H}}\boldsymbol{z}\rangle \tag{3.5.15}$$

となるが，これは $g^*: U\times V\to K$ が複素双線形形式であることを示している．g^* を定める線形写像を $A^{\mathrm{H}}: U\to V^*$(これを A に**反傾**な写像という)とすれば，A^{H} の表現行列は行列 A の **Hermite 共役** A^{H} である．

$g: V\times V\to K$ で $g=g^*$ のとき，複素双線形形式 g は **Hermite 的**(Hermitian)であるという．その表現行列は **Hermite 行列**である．

V の基底変換，U の基底変換(3.5.4)に伴って，複素双線形形式(3.5.14)の係数行列は

$$A'=S^{\mathrm{H}}AT \tag{3.5.16}$$

の形の変換を受ける．

$V=U$ のとき，基底変換も V と U で共通であるから，(3.5.16)は

$$A'=S^{\mathrm{H}}AS \tag{3.5.17}$$

という形になる(これも**合同変換**と呼ぶ)．S がユニタリ行列のときに限ってこの合同変換は相似変換に一致する．

3.5.3　二次形式，Hermite 形式

双線形形式 $g: V_n\times V_n\to K$ において $h(\boldsymbol{x})=g(\boldsymbol{x},\boldsymbol{x})$ とおくと $h: V_n\to K$ であるが，この h を**二次形式**(quadratic form)と呼ぶ．h を成分で表せば，g の係数行列 $A=[a_{ij}]$ を用いて，

$$h(\boldsymbol{x})=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x} \tag{3.5.18}$$

である．

注意 3.5.4　$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}$ を 1×1 行列とみなすと，$\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}=(\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x})^{\mathrm{T}}=\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x}$ であるから，

$$h(\boldsymbol{x})=\frac{1}{2}(\boldsymbol{x}^{\mathrm{T}}A\boldsymbol{x}+\boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\left(\frac{A+A^{\mathrm{T}}}{2}\right)\boldsymbol{x} \tag{3.5.19}$$

であり，「二次形式には係数行列の“対称部分”(注意 2.3.1)しか効かない」．したがって，二次形式について論じるときにはその係数行列は対称であると仮定するのが常である．双線形形式 $g: V_n\times V_n\to K$ についても，$A=\frac{1}{2}(A+A^{\mathrm{T}})+\frac{1}{2}(A-A^{\mathrm{T}})$ であるから，$g=g_{\mathrm{s}}+g_{\mathrm{a}}$ というように，g は対称な形式 g_{s} と交代形式 g_{a}

(→3.5.5 項)との和で表される．

以下，しばらく，$K=\mathbb{R}$ の場合を扱う．

対称行列の Cholesky 風の分解(2.15.4 項 a)を利用すれば，任意の二次形式 $h(\boldsymbol{x})=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_i x_j$ は，適当な基底 $\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$ を選ぶことにより

$$h(\boldsymbol{x})=x_1{}^2+\cdots+x_{r_+}{}^2-x_{r_++1}{}^2-\cdots-x_{r_++r_-}{}^2 \tag{3.5.20}$$

という形に表すことができる．具体的には，任意の基底に関する係数行列 A $(=A^{\mathrm{T}})$ に対して，(2.15.59)で求めた置換行列 P，下ブロック三角行列 L，対角行列 $D=\mathrm{diag}[d_1,d_2,\cdots,d_n]$ を用いた分解

$$PAP^{\mathrm{T}}=LDL^{\mathrm{T}} \tag{3.5.21}$$

は，$S=(L^{-1}P)^{\mathrm{T}}$ とおいた合同変換(3.5.7)により A が対角行列 D に変換されることを表している．そうした後，座標成分の順序を入れ換え，$\sqrt{|d_i|}$ ずつのスケール調整(これは各基底ベクトルを $\sqrt{|d_i|}$ 倍することに相当)すれば(3.5.20)の形に到達する．

注意 3.5.5 二次形式を(3.5.20)の“標準形”に変換する 2.15.4 項 a の方法は，古典的二次形式論で **Lagrange**(ラグランジュ)**の方法**と呼ばれていたものと本質的に同じである．

注意 3.5.6 V の上の二次形式は，その定義域を V の部分空間 U に制限することによって，U の上の二次形式を定める．

注意 3.5.7 (3.5.20)の形に二次形式を表してみれば，$r_++r_-=\mathrm{rank}\,D=\mathrm{rank}\,A$ であることは明らかである．二次形式を(3.5.20)の形にするような基底は一般に多数あるが，r_+,r_-(そして $r_0=n-(r_++r_-)$)はそのような基底の選び方によらない二次形式に固有の量である．この事実は **Sylvester の慣性則**(law of inertia)と呼ばれている．

順序づけられた三つ組 (r_+,r_-,r_0) $(r_++r_-=\mathrm{rank}\,A,\ r_++r_-+r_0=n)$ は A の **Sylvester の符号指数**(signature)と呼ばれる．(r_+-r_- のことを signature と呼び $\sigma(A)$ と記す慣習もあるが rank A と $\sigma(A)$ から r_+,r_- は定まり，また $r_0=n-(r_++r_-)$ から r_0 も定まるので，本質的には同じことである．) Sylvester の慣性則は，別のいい方をすると，「二つの正方対称行列が合同変換で結び合わされるための必要十分条件は，両者の符号指数が一致することである」ということを表している．

慣性則の証明 V の別の基底 $\{\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n'\}$ に関して

$$h(\boldsymbol{x})=x_1'^2+\cdots+x_{r_+'}'^2-x_{r_+'+1}'^2-\cdots-x_{r_+'+r_-'}'^2 \tag{3.5.22}$$

と表されたとする．いま，$r_+' \neq r_+$，たとえば $r_+' > r_+$ だったとし，V の部分空間 $U = L(\boldsymbol{e}_1', \cdots, \boldsymbol{e}'_{r_+'})$, $W = L(\boldsymbol{e}_{r_++1}, \cdots, \boldsymbol{e}_n)$ を作ると，(3.5.22)によれば任意の $\boldsymbol{x}$ $(\neq \boldsymbol{0}) \in U$ に対して $h(\boldsymbol{x}) > 0$，(3.5.20)によれば任意の $\boldsymbol{x}(\neq \boldsymbol{0}) \in W$ に対して $h(\boldsymbol{x}) \leqq 0$ であるから，$U \cap W = \{\boldsymbol{0}\}$．そこで直和空間 $U \oplus W$ を考えると，それは V の部分空間であるが，$\dim(U \oplus W) = \dim U + \dim W = r_+' + (n - r_+) > n$(→ 注意 3.2.5)．$\dim V = n$ であるから，このようなことは不可能である．したがって $r_+' \leqq r_+$．二つの基底の役割を入れ換えれば $r_+' \geqq r_+$ も示されるので，結局 $r_+' = r_+$ でなければならない．同様にして $r_-' = r_-$ でもある．[これから，既に示されている関係 $r'(= r_+' + r_-') = r(= r_+ + r_-)$ が再び得られる．] ∎

注意 3.5.8 Sylvester の慣性則の証明(注意 3.5.7)からも明らかなように，符号指数は

$r_+ = \max\{\dim U \mid$ “$\forall \boldsymbol{x}(\neq \boldsymbol{0}) \in U : h(\boldsymbol{x}) > 0$” であるような部分空間 $U\}$,
$r_- = \max\{\dim U \mid$ “$\forall \boldsymbol{x}(\neq \boldsymbol{0}) \in U : h(\boldsymbol{x}) < 0$” であるような部分空間 $U\}$,
$r_0 = \max\{\dim U \mid$ “$\forall \boldsymbol{x}(\neq \boldsymbol{0}) \in U : h(\boldsymbol{x}) = 0$” であるような部分空間 $U\}$

という特徴づけもできる．

$K = \mathbb{C}$ のとき，Hermite 的な複素双線形形式 g(すなわち $g : V_n \times V_n \to K$ で $g = g^*$ となるもの)に対して $h(\boldsymbol{x}) = g(\boldsymbol{x}, \boldsymbol{x})$ とおいたもの($h : V_n \to K$)を **Hermite(二次)形式**(Hermitian (quadratic) form)と呼ぶ．

$$h(\boldsymbol{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \bar{x}_i x_j \tag{3.5.23}$$

$g = g^*$ であるから，任意の $\boldsymbol{x} \in V_n$ に対して $\overline{h(\boldsymbol{x})} = h(\boldsymbol{x})$ であり，$h(\boldsymbol{x})$ は実数である．

Hermite 形式の議論は二次形式についての議論とほとんどすべて平行して進められる．たとえば，

- 係数行列は Hermite 行列と仮定してよい(→ 注意 3.5.10)；
- Hermite 行列に対する Cholesky 風の分解(2.15.60)を 2.15.4 項 b に従って修正したものを利用することによって，任意の Hermite 形式は V_n の適当な基底を選ぶことにより

$$h(x) = |x_1|^2 + \cdots + |x_{r_+}|^2 - |x_{r_++1}|^2 - \cdots - |x_{r_++r_-}|^2 \tag{3.5.24}$$

  という形に表すことができる；
- (r_+, r_-, r_0) $(r_0 = n - (r_+ + r_-))$ を係数行列 $A = [a_{ij}]$ の **Sylvester の符号指数**と呼ぶが，これは二次形式に固有の量で基底の選び方にはよらず(Sylvester

の慣性則)，注意 3.5.8 のような特徴づけもできる；
など．

注意 3.5.9　後に 5.2 節，6.5 節で述べるように，符号指数は係数行列の固有値(Hermite 行列に対しては実数である)の符号とも関係づけられる．しかし，合同変換は一般に固有値を不変に保たないこと，Cholesky 風分解は“有理演算”のみで遂行できるのに固有値計算は一般に“無理演算”(多項式の零点計算)を必要とすることに注意すべきである．一言でいえば，二次形式，Hermite 形式の話は第一義的には固有値とは無関係であるということである．

注意 3.5.10　Hermite 的でない複素双線形形式 g は，$A=\frac{1}{2}(A+A^{\mathrm{H}})+\mathrm{i}\frac{1}{2\mathrm{i}}(A-A^{\mathrm{H}})$ ($\frac{1}{2}(A+A^{\mathrm{H}})$ も $\frac{1}{2\mathrm{i}}(A-A^{\mathrm{H}})$ も Hermite 的)であるから，二つの Hermite 的な複素双線形形式 $g_{\mathrm{R}}, g_{\mathrm{I}}$ の和 $g=g_{\mathrm{R}}+\mathrm{i}g_{\mathrm{I}}$ として表される．

3.5.4　定符号形式，内積，計量

以下では $K=\mathbb{C}$ と仮定し Hermite 形式について述べる．$K=\mathbb{R}$ の場合の二次形式の話はその特殊な場合である．

任意の $\boldsymbol{x}(\neq\boldsymbol{0})\in V_n$ に対して $h(\boldsymbol{x})>0$ であるような Hermite 形式 $h: V_n\to\mathbb{R}$(あるいはその係数行列 $A(=A^*)$)は**正定値**(**正の定符号**：positive-definite)であるという．また，$h(\boldsymbol{x})\geqq 0$ のとき**非負定値**(“非”は“負定値”にではなく“負”だけにかかる；nonnegative-definite)あるいは**半正定値**(positive-semidefinite)という．同様にして，**負定値**(**負の定符号**；negative-definite)，**非正定値**(nonpositive-definite)，**半負定値**(negative-semidefinite)が定義される．

(半)負定値形式についての話は(半)正定値形式の符号を変えるだけのことであるから，以下では(半)正定値形式についてだけ述べる．

注意 3.5.11　正定値形式は符号指数が$(r_+, 0, 0)$であるもの，非負定値形式は符号指数が$(r_+, 0, r_0)$であるものと定義することができる．

注意 3.5.12　V_n の上の正定値形式は，V_n の任意の部分空間の上の正定値形式であることも，定義から直ちにわかる．

注意 3.5.13　n 次 Hermite 行列 $A(=A^{\mathrm{H}})$が正定値であるための必要十分条件の一つに，「そのすべての**首座小行列式**(→ 2.7.1 項)が正であること」：

$$|A|^{1\cdots k}_{1\cdots k}>0 \qquad (k=1,2,\cdots,n) \tag{3.5.25}$$

がある．[この事実は“理論的には”応用上も重要であるが，特定の行列の正定値性を計算で判定するには，Cholesky 分解(→ 2.15.4 項)を用いるのがよい．]

証明　A の k 次の首座小行列 $A^{1\cdots k}_{1\cdots k}$ も正定値であるから適当な合同変換により単位行列に変換できる（S は k 次正則）（→3.5.3 項）：

$$I_k = S^{\mathrm{H}} A^{1\cdots k}_{1\cdots k} S. \tag{3.5.26}$$

ゆえに，$1 = |\det S|^2 \cdot |A|^{1\cdots k}_{1\cdots k}$，すなわち $|A|^{1\cdots k}_{1\cdots k} > 0$．逆は n に関する帰納法による．$n=1$ のときは自明．ある $n(\geqq 2)$ に対して，(3.5.25) が成り立っているのに A が正定値ではなかったとする（帰納法の仮定により $A^{1,\cdots,n-1}_{1,\cdots,n-1}$ は正定値である）．ある n 次正則行列 S を用いた合同変換により A を対角行列 $D = \mathrm{diag}[d_1, \cdots, d_n] = S^{\mathrm{H}} A S$ に変換したとき，$|A|^{1\cdots n}_{1\cdots n} = \det A > 0$ より $\det D > 0$ であるので，A が，したがって D が，正定値でなければ少なくとも 2 個の d_i，たとえば d_{i_1} と d_{i_2} が負でなければならない．第 i_1（あるいは i_2）成分が 1 で他の成分がみな 0 の縦ベクトルを $\boldsymbol{e}_{i_1}$（あるいは $\boldsymbol{e}_{i_2}$）とし，$\boldsymbol{x} = \alpha \boldsymbol{e}_{i_1} + \beta \boldsymbol{e}_{i_2}$（$\alpha, \beta \in \mathbb{C}$；$\alpha = \beta = 0$ ではない）とおくと，$\boldsymbol{x}^{\mathrm{H}} D \boldsymbol{x} = |\alpha|^2 d_{i_1} + |\beta|^2 d_{i_2} = (S\boldsymbol{x})^{\mathrm{H}} A (S\boldsymbol{x}) < 0$．ところで，$\boldsymbol{y} = S\boldsymbol{x} = \alpha(S\boldsymbol{e}_{i_1}) + \beta(S\boldsymbol{e}_{i_2})$ の第 n 要素が 0 になるように α, β（$\alpha = \beta = 0$ ではない）を選べば（$y_n = \alpha S_{ni_1} + \beta S_{ni_2}$ であるからそのようなことはできる！），$\boldsymbol{y}^{\mathrm{H}} A \boldsymbol{y} < 0$ は $A^{1,\cdots,n-1}_{1,\cdots,n-1}$ が正定値であるという帰納法の仮定に反する．

注意 3.5.14　基底ベクトルの順序の入れ換え（A の行の置換とそれと同調した列の置換）をおこなってみれば明らかなように，「すべての**主小行列式**が正である」ことも A が正定値であるための必要かつ十分（すぎる）条件である．

$g: V_n \times V_n \to \mathbb{C}$ が Hermite 的な複素双線形形式であって，それから導かれる Hermite 形式 $h: V_n \to \mathbb{R}$ が正定値であるものを考える．g の係数行列 G は正定値 Hermite 行列で，V_n の基底を変換するとそれに伴って合同変換を受ける．双線形形式 g を，二つのベクトル $\boldsymbol{x}, \boldsymbol{y} (\in V_n)$ の間に一種の“積”を定義するものとみなして，（g, h あるいは G により定められる V_n の中の）**内積**(inner product) と呼び

$$(\!(\boldsymbol{x}, \boldsymbol{y})\!)_G = \langle \boldsymbol{x}, G\boldsymbol{y} \rangle \tag{3.5.27}$$

のように記す．そして，V_n のベクトルの“長さ”をノルム

$$\|\boldsymbol{x}\|_G = \sqrt{(\!(\boldsymbol{x}, \boldsymbol{x})\!)_G} \tag{3.5.28}$$

で定義する．このような V_n のことを**内積空間**あるいは G を**計量**(metric) とする**計量空間**(metric space) と呼ぶこともある．

注意 3.5.15　(3.5.28) の $\|\boldsymbol{x}\|_G$ が実際にノルムの条件（3.1.5 項の (1)，(2)，(3)）を満たすことは，V_n の基底として $G = I_n$ となるようなものを選ぶと，$\|\boldsymbol{x}\|_G = \|\boldsymbol{x}\|_2$（→ 3.1.5 項）となることから明らかである．

ここで注意すべきことは，p 乗ノルムはベクトルに固有の量ではなくて基底に依存する量であるが，内積空間におけるノルムはベクトルに固有の量で，基底が異なってもベクトルの座標成分によるその表現(行列 G)が異なるだけでこのノルムの値は不変であるということである．応用分野ごとに適切な計量を定めてベクトルの長さを計るようにすることが大切で，たまたま用いている基底(座標系)に関して無分別に 2 乗ノルムなどを使うようなことはすべきでない．

注意 3.5.16 二つの Hermite 形式 h, h'(あるいは Hermite 行列 A, A')の差 $h-h'$(あるいは $A-A'$)も Hermite 形式(行列)であるが，それが非負定値であるとき

$$h \geqq h' \quad \text{あるいは} \quad A \geqq A' \tag{3.5.29}$$

と表すことにすると，$\geqq$ は同じベクトル空間 V_n の上の Hermite 形式(同一の次数の Hermite 行列)の間の一つの半順序(partial order)になっている(一般には全順序にはならない)．二つの非負定値 Hermite 形式(行列)の和や正の実数倍も非負定値である．

3.5.5 交代形式

2.15.4 項 c で述べた反対称行列の Cholesky 風分解(2.15.65)を双線形形式の言葉で述べ直すことができる．

双線形形式 $g: V_n \times V_n \to K$ は $g(\boldsymbol{x}, \boldsymbol{y}) = -g(\boldsymbol{y}, \boldsymbol{x})$ $(\forall \boldsymbol{x}, \boldsymbol{y} \in V_n)$のとき**交代形式**(alternating form)であるといわれる．g を定める線形写像を $A: V_n \to V_n^{\mathrm{d}}$ とすれば，

$$g(\boldsymbol{x}, \boldsymbol{y}) = \langle \boldsymbol{y}, A\boldsymbol{x} \rangle = \langle \boldsymbol{x}, A^{\top}\boldsymbol{y} \rangle = -g(\boldsymbol{y}, \boldsymbol{x}) = -\langle \boldsymbol{x}, A\boldsymbol{y} \rangle \tag{3.5.30}$$

である．任意の $\boldsymbol{x}, \boldsymbol{y}$ に対して(3.5.30)が成り立つのであるから，$A^{\top} = -A$ であり，写像(そしてその表現行列)A は**反対称**(anti-symmetric)である．V_n の適当な基底を選ぶと，rank $A = 2p$ のとき，

$$g(\boldsymbol{x}, \boldsymbol{y}) = (x_1 y_2 - x_2 y_1) + (x_3 y_4 - x_4 y_3) + \cdots + (x_{2p-1} y_{2p} - x_{2p} y_{2p-1}) \tag{3.5.31}$$

という形に表せる．

注意 3.5.17 3.5.5 項の話はこれだけであれば K は任意の体でよい．

3.6 正規直交系

V_n が G を計量とする $K=\mathbb{C}$ の上の計量空間であるとする．ベクトル $\boldsymbol{x}, \boldsymbol{y}$ $(\in V_n)$ は $\langle\!\langle \boldsymbol{x}, \boldsymbol{y} \rangle\!\rangle_G (=\overline{\langle\!\langle \boldsymbol{y}, \boldsymbol{x} \rangle\!\rangle_G})=0$ のとき，すなわち内積が0であるとき，互いに**直交**する(to be orthogonal to each other)という．また，$\boldsymbol{x}$ は $\|\boldsymbol{x}\|_G{}^2=\langle\!\langle \boldsymbol{x}, \boldsymbol{x} \rangle\!\rangle_G=1$ であるとき**単位長ベクトル**(unit-length vector)という(単に**単位ベクトル**(unit vector)ということもある)．

ベクトルの集合$\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_r\}$は

$$\left.\begin{aligned} &\|\boldsymbol{x}_i\|_G=1 \qquad (i=1,\cdots,r), \\ &\langle\!\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle\!\rangle_G=0 \qquad (i\neq j\ ;\ i,j=1,\cdots,r) \end{aligned}\right\} \tag{3.6.1}$$

であるとき**正規直交系**(orthonormal system)をなすといわれる．

注意 3.6.1 正規直交系は線形独立である．実際

$$\sum_{j=1}^{r} \lambda_j \boldsymbol{x}_j = \boldsymbol{0} \tag{3.6.2}$$

なら，両辺の $\boldsymbol{x}_i$ との内積を取って，

$$\sum_{j=1}^{r} \lambda_j \langle\!\langle \boldsymbol{x}_i, \boldsymbol{x}_j \rangle\!\rangle_G = \lambda_i = 0 \tag{3.6.3}$$

であるからである．

正規直交系$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$が V_n の基底であるとき，特に，**正規直交基底**(orthonormal basis)という．

一般のベクトルの集合$\{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_m\}$$(\boldsymbol{y}_i \in V_n)$が与えられたとき，$\{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_m\}$の張る部分空間 $L(\boldsymbol{y}_1, \cdots, \boldsymbol{y}_m)$の正規直交基底を具体的に構成する手続きとしてよく知られているのが以下に述べる **Gram-Schmidt**(グラム-シュミット)**の直交化法**(Gram-Schmidt orthogonalization)である(**Schmidt の直交化法**と呼ぶこともある)．

[1]　$i:=0$，$k:=1$，$\boldsymbol{z}_1:=\boldsymbol{y}_1$ とおいて[2]へ進む．

[2]

　[2.1]　$\boldsymbol{z}_k=\boldsymbol{0}$ なら[3]へ進む．

[2.2] $\boldsymbol{z}_k \neq \boldsymbol{0}$ なら $i := i+1$ とおいて

$$\boldsymbol{e}_i := \boldsymbol{z}_k / \|\boldsymbol{z}_k\|_G \tag{3.6.4}$$

を作り，[3]へ進む．

[3]

[3.1] $k=m$ なら終了．

[3.2] $k<m$ なら

$$\boldsymbol{z}_{k+1} := \boldsymbol{y}_{k+1} - \sum_{j=1}^{i} \langle\!\langle \boldsymbol{e}_j, \boldsymbol{y}_{k+1} \rangle\!\rangle_G \, \boldsymbol{e}_j \tag{3.6.5}$$

を作り，$k := k+1$ とおいて[2]へ戻る．

上記の算法で所望の基底$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_r\}$が得られることは次のようにして示される．[3.1]で算法が終了したときの i の値を r とおく．まず，$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_i\}$が $i=1, \cdots, r$ に対しすべて正規直交系をなすことを示す．[(3.6.4)により $\|\boldsymbol{e}_i\|_G = 1$ である．ある i に対して$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_i\}$が正規直交系，すなわち$\langle\!\langle \boldsymbol{e}_j, \boldsymbol{e}_l \rangle\!\rangle_G = 0$ $(j, l \leqq i\,;\, j \neq l)$であるとすれば，(3.6.5)により $l=1, \cdots, i$ に対して$\langle\!\langle \boldsymbol{e}_l, \boldsymbol{z}_{k+1} \rangle\!\rangle_G = \langle\!\langle \boldsymbol{e}_l, \boldsymbol{y}_{k+1} \rangle\!\rangle_G - \sum_{j=1}^{i} \langle\!\langle \boldsymbol{e}_j, \boldsymbol{y}_{k+1} \rangle\!\rangle_G \cdot \langle\!\langle \boldsymbol{e}_l, \boldsymbol{e}_j \rangle\!\rangle_G = \langle\!\langle \boldsymbol{e}_l, \boldsymbol{y}_{k+1} \rangle\!\rangle_G - \langle\!\langle \boldsymbol{e}_l, \boldsymbol{y}_{k+1} \rangle\!\rangle_G = 0$ であるから$\langle\!\langle \boldsymbol{e}_l, \boldsymbol{e}_{i+1} \rangle\!\rangle_G = 0$ $(l=1, \cdots, i)$で，$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_{i+1}\}$も正規直交系である．]次に，(3.6.4)において対応する i と k に対して，$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_i\}$が $L(\boldsymbol{y}_1, \cdots, \boldsymbol{y}_k)$の基底になっていることを示せばよい．[(3.6.5)で $\boldsymbol{z}_{k+1} = \boldsymbol{0}$ なら $\boldsymbol{y}_{k+1}$ はそれまでに作られた$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_i\}$に従属．$\boldsymbol{z}_{k+1} \neq \boldsymbol{0}$ なら次に作られる $\boldsymbol{e}_{i+1}$ を用いて $\boldsymbol{z}_{k+1} = \|\boldsymbol{z}_{k+1}\|_G \, \boldsymbol{e}_{i+1}$ と表せるから $\boldsymbol{y}_{k+1}$ は$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_{i+1}\}$に従属．]

注意 3.6.2 上の算法およびその正当性の証明において既に用い証明した事実ではあるが，$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_r\}$が V_n の部分空間 U の基底であるとき，任意の $\boldsymbol{y}(\in V_n)$を $\boldsymbol{y} = \boldsymbol{y}_{/\!/} + \boldsymbol{y}_{\perp}$ $(\boldsymbol{y}_{/\!/} \in U\,;$ すべての $\boldsymbol{x}(\in U)$に対して$\langle\!\langle \boldsymbol{y}_{\perp}, \boldsymbol{x} \rangle\!\rangle_G = 0)$と分けることができる．$\boldsymbol{y}_{/\!/}$ と $\boldsymbol{y}_{\perp}$ は

$$\boldsymbol{y}_{/\!/} = \sum_{j=1}^{r} \langle\!\langle \boldsymbol{e}_j, \boldsymbol{y} \rangle\!\rangle_G \, \boldsymbol{e}_j, \qquad \boldsymbol{y}_{\perp} = \boldsymbol{y} - \boldsymbol{y}_{/\!/} \tag{3.6.6}$$

により定められる．$\boldsymbol{y}_{/\!/}$ は $\boldsymbol{y}$ の U への**直交射影**(orthogonal projection)と呼ばれる．

V_n の基底として$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_r, \boldsymbol{e}_{r+1}, \cdots, \boldsymbol{e}_n\}$の形のものを取り，各 $\boldsymbol{e}_{r+k}$ から$(\boldsymbol{e}_{r+k})_{\perp}$を(3.6.6)により作れば$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_r, (\boldsymbol{e}_{r+1})_{\perp}, \cdots, (\boldsymbol{e}_n)_{\perp}\}$も V_n の基底(正規直交であるとは限らない)である．$W = L((\boldsymbol{e}_{r+1})_{\perp}, \cdots, (\boldsymbol{e}_n)_{\perp})$は U の補空間で，$\forall \boldsymbol{x} \in U, \forall \boldsymbol{y}$

$\in W$ に対して$\langle\!\langle \boldsymbol{x}, \boldsymbol{y} \rangle\!\rangle_G=0$ となる．すなわち W は U の**直交補空間**(orthogonal complement)である．U の補空間は一意に決まらないが，直交補空間(それを $U^\perp$ と記す)は一意に定まる．

注意 3.6.3 Gram-Schmidt の直交化法には次のような変種もあり，数値計算上はこの方が丸め誤差の悪影響を受け難く，すぐれているという説もあるが，詳しくは線形計算の専門書を参照されたい．(一言でいえば，(3.6.5)において $\boldsymbol{y}_{k+1}$ から差し引く $\boldsymbol{e}_j$ に比例したベクトルをあらかじめ引けるようになったら早々と引いてしまうというやり方である．)

[1′] $i:=0, k:=1$ とおいて[2′]へ進む．

[2′]

[2.1′] $\boldsymbol{y}_k=\boldsymbol{0}$ なら $k:=k+1$ とおく．$k \leqq m$ なら[2′]に戻り，$k=m+1$ なら終了．

[2.2′] $\boldsymbol{y}_k \neq \boldsymbol{0}$ なら $i:=i+1$ とおいて

$$\boldsymbol{e}_i := \boldsymbol{y}_k / \|\boldsymbol{y}_k\|_G \tag{3.6.7}$$

を作り，[3′]へ進む．

[3′]

[3.1′] $k=m$ なら終了．

[3.2′] $k<m$ なら $l:=k+1,\cdots,m$ に対して

$$\boldsymbol{y}_l := \boldsymbol{y}_l - \langle\!\langle \boldsymbol{e}_i, \boldsymbol{y}_l \rangle\!\rangle_G\, \boldsymbol{e}_i \tag{3.6.8}$$

を作り，$k:=k+1$ とおいて[2′]へ戻る．

注意 3.6.4 $B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$，$B'=\{\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n'\}$が共に正規直交系であるとき，その間の変換式

$$[\boldsymbol{e}_1',\cdots,\boldsymbol{e}_n'] = [\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n]Q \tag{3.6.9}$$

に現れる行列 $Q=[q_{ij}]$ は，

$$\begin{aligned}\langle\!\langle \boldsymbol{e}_i', \boldsymbol{e}_j' \rangle\!\rangle_G &= \left\langle\!\!\left\langle \sum_{k=1}^n q_{ki}\boldsymbol{e}_k, \sum_{l=1}^n q_{lj}\boldsymbol{e}_l \right\rangle\!\!\right\rangle_G \\ &= \sum_{k=1}^n \sum_{l=1}^n \bar{q}_{ki} q_{lj} \langle\!\langle \boldsymbol{e}_k, \boldsymbol{e}_l \rangle\!\rangle_G = \sum_{k=1}^n \bar{q}_{ki} q_{kj}\end{aligned} \tag{3.6.10}$$

と$\langle\!\langle \boldsymbol{e}_i', \boldsymbol{e}_j' \rangle\!\rangle_G=\delta_{ij}$ とにより，$Q^{\mathrm{H}}Q=I_n$ という条件を満たさなければならない．すなわち Q はユニタリ行列でなければならない．逆に，B が正規直交で Q がユニタリなら(3.6.9)で定められる B' は正規直交である．

注意 3.6.5 G を計量とする計量空間 V_n では $G: V_n \to V_n^*$ とみなせるが，$B=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$が V_n の正規直交基底であれば，$\boldsymbol{e}_i^*=G\boldsymbol{e}_i$ により定められる V_n^* の基

底 $B^*=\{\boldsymbol{e}_1^*,\cdots,\boldsymbol{e}_n^*\}$ は B に対応する双対基底である．このことは $\langle\!\langle \boldsymbol{e}_i,\boldsymbol{e}_j\rangle\!\rangle_G=\langle \boldsymbol{e}_i,G\boldsymbol{e}_j\rangle=\langle \boldsymbol{e}_i,\boldsymbol{e}_j^*\rangle(=\delta_{ij})$ という定義関係式の読み換えにすぎない．

注意 3.6.6 $\{\boldsymbol{x}_1,\cdots,\boldsymbol{x}_m\}$ が $\langle\!\langle \boldsymbol{x}_i,\boldsymbol{x}_j\rangle\!\rangle_G=0\,(i\neq j)$，$\|\boldsymbol{x}_i\|\neq 0$ を満たすとき，**直交系** (orthogonal system) と呼ぶ．注意 3.6.1 で示したのと同様にして，直交系は線形独立であることが示せる．また，直交系に対しては

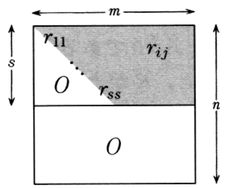
$$\|\boldsymbol{x}_1+\cdots+\boldsymbol{x}_m\|_G{}^2=\left\langle\!\!\left\langle \sum_{i=1}^m \boldsymbol{x}_i,\sum_{j=1}^m \boldsymbol{x}_j\right\rangle\!\!\right\rangle_G=\|\boldsymbol{x}_1\|_G{}^2+\cdots+\|\boldsymbol{x}_m\|_G{}^2 \qquad (3.6.11)$$

すなわち **Pythagoras**（ピュタゴラス）**の定理**が成り立つ．

$L(\boldsymbol{y}_1,\cdots,\boldsymbol{y}_m)$ の正規直交基底を構成するには，行列の QR 分解（→ 2.15.5 項）を利用することもできる．話を簡単にするために，内積の係数行列 G が I_n になるように V_n の基底（正規直交基底；注意 3.6.7 を見よ）を適当に選んでおく．その基底に関する $\boldsymbol{y}_i$ の座標成分を行列の形に並べて $n\times m$ 行列 Y とし，その QR 分解（の複素版）

$$YP=[\boldsymbol{y}_1,\cdots,\boldsymbol{y}_m]P=QR \qquad (3.6.12)$$

を作る．ここで，P は置換行列（$\boldsymbol{y}_1,\cdots,\boldsymbol{y}_m$ の順序の入れ換えにすぎない），Q は n 次のユニタリ行列，R は上三角行列

$R=$ 

$(r_{ii}\neq 0;\ i=1,\cdots,s)$ (3.6.13)

である．Q の列ベクトルを $\boldsymbol{q}_{\cdot 1},\cdots,\boldsymbol{q}_{\cdot n}$ とすると，$\langle\!\langle \boldsymbol{q}_{\cdot i},\boldsymbol{q}_{\cdot j}\rangle\!\rangle_G=\delta_{ij}$ であるから $\{\boldsymbol{q}_{\cdot 1},\cdots,\boldsymbol{q}_{\cdot n}\}$ が V_n の正規直交基底であり，(3.6.12) は（$\boldsymbol{y}_1,\cdots,\boldsymbol{y}_m$ に P が指定する順序入れ換えを施したものを $\tilde{\boldsymbol{y}}_1,\cdots,\tilde{\boldsymbol{y}}_m$ として）

$$\left.\begin{aligned}\tilde{\boldsymbol{y}}_j&=\sum_{i=1}^{j} r_{ij}\boldsymbol{q}_{\cdot i} &&(j=1,\cdots,s),\\ \tilde{\boldsymbol{y}}_j&=\sum_{i=1}^{s} r_{ij}\boldsymbol{q}_{\cdot i} &&(j=s+1,\cdots,m)\end{aligned}\right\} \qquad (3.6.14)$$

であることを表している．これは正規直交系 $\{\boldsymbol{q}_{\cdot 1},\cdots,\boldsymbol{q}_{\cdot s}\}$ が $L(\boldsymbol{y}_1,\cdots,\boldsymbol{y}_m)$ の基底であることを示している．［$\boldsymbol{q}_{\cdot i}\in L(\boldsymbol{y}_1,\cdots,\boldsymbol{y}_m)\ (i=1,\cdots,s)$ であることは

$$[\tilde{\boldsymbol{y}}_1, \cdots, \tilde{\boldsymbol{y}}_s] = [\boldsymbol{q}_{\cdot 1}, \cdots, \boldsymbol{q}_{\cdot s}] \cdot \begin{bmatrix} r_{11} & & * \\ & \ddots & \\ O & & r_{ss} \end{bmatrix} \tag{3.6.15}$$

を逆に解いて，$\boldsymbol{q}_{\cdot i}$ が $\tilde{\boldsymbol{y}}_j$ の線形結合で表せることから明らか.]

要するに $L(\boldsymbol{y}_1, \cdots, \boldsymbol{y}_m)$ の正規直交基底を作る話と行列の QR 分解とは実質的に同じであるということである.

注意 3.6.7　V_n 自身の正規直交基底を $B=\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$ とし，B に関するベクトル $\boldsymbol{x}, \boldsymbol{y}$ の座標成分を x_i, y_i とすれば，

$$\langle\!\langle \boldsymbol{x}, \boldsymbol{y} \rangle\!\rangle_G = \sum_{i=1}^{n} \sum_{j=1}^{n} \bar{x}_i y_j \langle\!\langle \boldsymbol{e}_i, \boldsymbol{e}_j \rangle\!\rangle_G = \sum_{i=1}^{n} \bar{x}_i y_i \tag{3.6.16}$$

となるから，B に関しては計量 G の係数行列は単位行列となる．すなわち，G を合同変換により単位行列に変換することと正規直交基底を作ることとは実質的に同じ話である.

注意 3.6.8　n 次正方行列 $A=[a_{ij}]$ の QR 分解の副産物として，$\det A$ に対する **Hadamard の不等式**(Hadamard's inequality)

$$|\det A| \leqq \prod_{j=1}^{n} \left(\sum_{i=1}^{n} |a_{ij}|^2 \right)^{1/2} \tag{3.6.17}$$

が得られる．$\operatorname{rank} A < n$ のときには $\det A = 0$ であるから(3.6.17)は成り立つ．$\operatorname{rank} A = n$ のときには，$A=QR$ とすると，R が上三角で正則である．このとき，A の列ベクトルを $\boldsymbol{a}_{\cdot j}$, Q の列ベクトルを $\boldsymbol{q}_{\cdot i}$ として，

$$\boldsymbol{a}_{\cdot j} = \sum_{i=1}^{j} r_{ij} \boldsymbol{q}_{\cdot i} \qquad (j=1, \cdots, n) \tag{3.6.18}$$

と表せるが，Pythagoras の定理(注意 3.6.6)により，

$$\|\boldsymbol{a}_{\cdot j}\|_2{}^2 = \sum_{i=1}^{n} |a_{ij}|^2 = \sum_{i=1}^{j} |r_{ij}|^2 \|\boldsymbol{q}_{\cdot i}\|_2{}^2 = \sum_{i=1}^{j} |r_{ij}|^2 \geqq |r_{jj}|^2. \tag{3.6.19}$$

また，$A=QR$，$|\det Q|=1$ より，

$$|\det A| = |\det R| = \prod_{j=1}^{n} |r_{jj}| \tag{3.6.20}$$

が成り立つ．(3.6.19)，(3.6.20)から(3.6.17)が導かれる.

3.7　射　　影

3.7.1　射　　影

n 次元ベクトル空間 V_n の直和分解 $V_n = U \oplus W$ (→ 3.2.2 項)が与えられたと

き，$\boldsymbol{x}(\in V_n)$に対して，$\boldsymbol{x}=\boldsymbol{x}_U+\boldsymbol{x}_W(\boldsymbol{x}_U\in U, \boldsymbol{x}_W\in W)$と(一意的に)表すときの，$\boldsymbol{x}_U$ を対応させる線形写像 $P_{U,W}: \boldsymbol{x}\mapsto \boldsymbol{x}_U$ を“W に沿っての U の上への”**射影**(projection)と呼ぶ．$P_{W,U}: \boldsymbol{x}\mapsto \boldsymbol{x}_W$ は U に沿っての W の上への射影であって，$I=P_{U,W}+P_{W,U}$(=恒等写像)である．

U に基底 $B_U=\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_s\}$，$W$ に基底 $B_W=\{\boldsymbol{e}_{s+1},\cdots,\boldsymbol{e}_n\}$を選べば $B=B_U\cup B_W$ は V_n の基底である．B の双対基底 $B^{\mathrm{d}}=\{\boldsymbol{e}_1^{\mathrm{d}},\cdots,\boldsymbol{e}_n^{\mathrm{d}}\}$から $B_U^{\mathrm{d}}=\{\boldsymbol{e}_1^{\mathrm{d}},\cdots,\boldsymbol{e}_s^{\mathrm{d}}\}$，$B_W^{\mathrm{d}}=\{\boldsymbol{e}_{s+1}^{\mathrm{d}},\cdots,\boldsymbol{e}_n^{\mathrm{d}}\}$を作れば，$U^{\mathrm{d}}=L(B_U^{\mathrm{d}}), W^{\mathrm{d}}=L(B_W^{\mathrm{d}}), V_n^{\mathrm{d}}=U^{\mathrm{d}}\oplus W^{\mathrm{d}}$ である．(一般には，部分空間 U の双対空間は V_n の双対空間 V_n^{d} の部分空間ではないが，ここではそれが V_n^{d} の部分空間とみなせるように補空間 W も含めて定義していることに注意！）これらを用いると，$P_{U,W}, P_{W,U}$ はより具体的に

$$P_{U,W}\,\boldsymbol{x}=\boldsymbol{x}_U=\sum_{i=1}^{s}\langle\boldsymbol{x},\boldsymbol{e}_i^{\mathrm{d}}\rangle\boldsymbol{e}_i,\qquad P_{W,U}\,\boldsymbol{x}=\boldsymbol{x}_W=\sum_{i=s+1}^{n}\langle\boldsymbol{x},\boldsymbol{e}_i^{\mathrm{d}}\rangle\boldsymbol{e}_i \tag{3.7.1}$$

と表せる．基底 B に関する $P_{U,W}$ の表現行列は

$$P_{U,W}=\left[\begin{array}{c|c} \begin{matrix}1 & & O\\ & \ddots & \\ O & & 1\end{matrix} & O \\ \hline O & O\end{array}\right]\begin{matrix} \}\,s \\ \}\,n-s\end{matrix} \quad (\text{列: } s,\ n-s) \tag{3.7.2}$$

である．$P=P_{U,W}$ は，明らかに，

$$P^2=P \tag{3.7.3}$$

を満たす．

一般に，(3.7.3)を満たす線形写像 $P: V_n\to V_n$ のことを**射影**と呼ぶ．実際，$U=\mathrm{Im}\,P, W=\mathrm{Ker}\,P$ とおくと，$\boldsymbol{x}\in U\cap W$ なら，$\boldsymbol{x}\in U$ なので $\boldsymbol{x}=P\boldsymbol{y}$ となるような $\boldsymbol{y}(\in V_n)$が存在するが，他方 $\boldsymbol{x}\in W$ より $P\boldsymbol{x}=\boldsymbol{0}$．そこで(3.7.3)により $P\boldsymbol{x}=P^2\boldsymbol{y}=P\boldsymbol{y}=\boldsymbol{x}=\boldsymbol{0}$ となってしまう．このことは $U\cap W=\{\boldsymbol{0}\}$，すなわち，$U+W$ が直和 $U\oplus W$ であることを示している．ところが，(3.3.34)より $\dim U+\dim W=n$ であるから $V_n=U\oplus W$．したがって，P が(3.7.3)を満たせば $P=P_{\mathrm{Im}\,P,\mathrm{Ker}\,P}$．

注意 3.7.1　一般に(3.7.3)を満たす写像 $P: V_n\to V_n$ あるいは正方行列 P のことを**べき等**(idempotent)であるという．もちろん，P がべき等なら $P^k=P\,(k\geqq 1)$ である．

P が射影のとき $P'=I-P$ も射影で $P'=P_{\mathrm{Ker}\,P,\mathrm{Im}\,P}$ である．実際，$P'^2=(I-P)\cdot(I-P)=I-2P+P^2=I-P=P'$ である．さらに，$\boldsymbol{x}\in\mathrm{Ker}\,P'$ なら，$P'\boldsymbol{x}=\boldsymbol{x}-P\boldsymbol{x}=\boldsymbol{0}$，すなわち $\boldsymbol{x}=P\boldsymbol{x}$，ゆえに $\boldsymbol{x}\in\mathrm{Im}\,P$．また，$\boldsymbol{x}\in\mathrm{Im}\,P'$ なら，$\boldsymbol{x}=P'\boldsymbol{y}=\boldsymbol{y}-P\boldsymbol{y}$ より $P\boldsymbol{x}=P\boldsymbol{y}-P^2\boldsymbol{y}=P\boldsymbol{y}-P\boldsymbol{y}=\boldsymbol{0}$ で，$\boldsymbol{x}\in\mathrm{Ker}\,P$．

また，P と $P'=I-P$ とは明らかに**互いに消し合う**，すなわち

$$PP'=P'P=O \tag{3.7.4}$$

という関係にある．そして，$\boldsymbol{x}=\boldsymbol{x}_U+\boldsymbol{x}_W$ ($\boldsymbol{x}\in V_n$，$\boldsymbol{x}_U\in U$，$\boldsymbol{x}_W\in W$) は

$$\boldsymbol{x}_U=P\boldsymbol{x}, \qquad \boldsymbol{x}_W=P'\boldsymbol{x} \tag{3.7.5}$$

と表せる．

注意 3.7.2　P が射影なら，V_n の適当な基底を用いて，(3.7.2) の形の行列表現をもつ．このとき $\dim(\mathrm{Im}\,P)=s=\mathrm{tr}\,P$ (→ 2.7.2 項)．一般に，行列のトレースは相似変換によって変化しない[(2.7.4) により $\mathrm{tr}(S^{-1}AS)=\mathrm{tr}(ASS^{-1})=\mathrm{tr}\,A$] から，射影 P の任意の基底に関する表現行列のトレースは $\dim(\mathrm{Im}\,P)$ に等しい．

一般に，V_n が 3 個以上の部分空間の直和

$$V_n=U_1\oplus\cdots\oplus U_m \tag{3.7.6}$$

であるときも，U_i の補空間を $W_i=U_1\oplus\cdots\oplus U_{i-1}\oplus U_{i+1}\oplus\cdots\oplus U_m$ として，$P_i=P_{U_i,W_i}$ とすれば $P_1,\cdots,P_m$ は射影で

$$P_i^2=P_i, \tag{3.7.7a}$$

$$I=P_1+\cdots+P_m, \tag{3.7.7b}$$

$$P_iP_j=O \qquad (i\neq j) \tag{3.7.7c}$$

を満たす．逆に，(3.7.7) を満たす線形写像 $P_1,\cdots,P_m$ があるとき，$U_i=\mathrm{Im}\,P_i$ とおけば P_i は (3.7.7a) により U_i の上への射影であり，(3.7.7c) により注意 3.2.5 (iv) の条件が満たされるから，$U_1+\cdots+U_m$ は直和 $U_1\oplus\cdots\oplus U_m$ である．そして (3.7.7b) により $V_n=U_1\oplus\cdots\oplus U_m$ である．

注意 3.7.3　(3.7.7) の条件はすべて独立ではない．実際，線形写像 $A_i: V_n\to V_n$ ($i=1,\cdots,m$) が (3.7.7b) の条件

$$I=A_1+\cdots+A_m \tag{3.7.8}$$

を満たすとき，$U_i=\mathrm{Im}\,A_i$ とおけば，次の 4 条件は互いに等価である．

(i)　$V_n=U_1\oplus\cdots\oplus U_m$.　(3.7.9)

(ii)　$\dim U_1+\cdots+\dim U_m\leqq n$.　(3.7.10)

(iii)　$A_iA_j=O$　　$(i\neq j\,;\ i,j=1,\cdots,m)$.　(3.7.11)

(iv)　$A_i^2=A_i$　　$(i=1,\cdots,m)$.　(3.7.12)

略証　(ii)$\Longrightarrow$(i)：$U_i\subseteqq V_n$ だから $U_1+\cdots+U_m\subseteqq V_n$. 任意の $\boldsymbol{x}(\in V_n)$ に対し(3.7.8)より $\boldsymbol{x}=A_1\boldsymbol{x}+\cdots+A_m\boldsymbol{x}$ だから $V_n\subseteqq U_1+\cdots+U_m$. そこで $V_n=U_1+\cdots+U_m$. したがって $n\leqq \dim U_1+\cdots+\dim U_m$. (ii)と合わせれば $n=\dim U_1+\cdots+\dim U_m$. 注意3.2.5により(i)が得られる．(i)$\Longrightarrow$(iii)：$\boldsymbol{x}(\in V_n)$に対し $A_j\boldsymbol{x}\in U_j$. (3.7.8)により $IA_j\boldsymbol{x}=A_j\boldsymbol{x}=A_1A_j\boldsymbol{x}+\cdots+A_jA_j\boldsymbol{x}+\cdots+A_mA_j\boldsymbol{x}$. $A_i(A_j\boldsymbol{x})\in U_i$ であるから(i)(分解の一意性)により $A_iA_j\boldsymbol{x}=\boldsymbol{0}\,(i\neq j)$, $A_j^2\boldsymbol{x}=A_j\boldsymbol{x}$. すなわち(iii)を(そして(iv)も)得る．(iii)$\Longrightarrow$(iv)：(3.7.8)の両辺に A_i を掛けると，(iii)により $A_i=A_i^2$. (iv)$\Longrightarrow$(ii)：V_n に任意の基底を選んで A_i を行列表現しておく．(iv)により A_i は射影であるから，注意3.7.2により $\dim U_i=\mathrm{tr}\,A_i$. (3.7.8)により $\mathrm{tr}\,I=n=\mathrm{tr}\,A_1+\cdots+\mathrm{tr}\,A_m=\dim U_1+\cdots+\dim U_m$. すなわち(ii)を得る．■

3.7.2 正　射　影

$K=\mathbb{C}$ の上のベクトル空間 V_n が G を計量とする計量空間であるとする．射影 $P=P_{\mathrm{Im}\,P,\mathrm{Ker}\,P}$ において $\mathrm{Im}\,P$ と $\mathrm{Ker}\,P$ とが互いに直交補空間であるとき，P を**正射影**(orthogonal projection)と呼ぶ(→ 注意3.6.2)．このとき，任意の $\boldsymbol{x},\boldsymbol{y}\in V_n$ に対して，

$$\langle\!\langle (I-P)\boldsymbol{x},P\boldsymbol{y}\rangle\!\rangle_G=((I-P)\boldsymbol{x})^{\mathrm{H}}GP\boldsymbol{y}=0, \tag{3.7.13}$$

すなわち

$$\boldsymbol{x}^{\mathrm{H}}(GP-P^{\mathrm{H}}GP)\boldsymbol{y}=0 \tag{3.7.14}$$

であるから

$$GP-P^{\mathrm{H}}GP=O \quad \text{あるいは} \quad (G^{-1}P^{\mathrm{H}}G)P=P \tag{3.7.15}$$

である．以後，G を計量とする計量空間 V_n における線形写像 A に対して

$$A^{\#}=G^{-1}A^{\mathrm{H}}G \tag{3.7.16}$$

と，ときどき書くことがある．(A^{H} は写像 $A：V_n\to V_n$ の表現行列の Hermite 共役を表現行列とする写像で，$A^{\mathrm{H}}：V_n^*\to V_n^*$ とみなす；3.5.2項において $U=V^*$ としたもの；正規直交基底に関する行列表現を用いれば $G=I$ であるから $A^{\#}=A^{\mathrm{H}}$.) すると，(3.7.15)の条件は

$$P^{\#}P=P \tag{3.7.17}$$

と表せる．

注意 3.7.4 実は(3.7.17)の条件は P が射影であるという条件(3.7.3)を含んでいる．実際，(3.7.17)が成り立てば

$$P^{\#}=(P^{\#}P)^{\#}=G^{-1}(G^{-1}P^{\mathrm{H}}GP)^{\mathrm{H}}G=G^{-1}P^{\mathrm{H}}G^{\mathrm{H}}PG^{\mathrm{H}-1}G.$$

$G^{\mathrm{H}}=G$ であるから，

$$P^{\#}=G^{-1}P^{\mathrm{H}}GP=P. \tag{3.7.18}$$

(3.7.18)，(3.7.17)より(3.7.3)が導かれる．また，P が射影で(3.7.18)を満たせば(3.7.17)を満たすので，P は正射影である．

注意 3.7.5 $\boldsymbol{y}\in V_n$ の U への正射影 $\boldsymbol{x}=P\boldsymbol{y}$ は，$\boldsymbol{y}$ との“距離”$\|\boldsymbol{y}-\boldsymbol{x}\|_G$ が最小になるような U の中のベクトル $\boldsymbol{x}$ であるという特徴づけもできる．実際，$\|\boldsymbol{y}-\boldsymbol{x}\|_G$ を最小にする U のベクトルの一つを $\hat{\boldsymbol{x}}$ とし，$\boldsymbol{z}$ を U の任意のベクトル，λ を任意の複素数とすると，

$$\begin{aligned}\|\boldsymbol{y}-(\hat{\boldsymbol{x}}+\lambda\boldsymbol{z})\|_G{}^2&=\|\boldsymbol{y}-\hat{\boldsymbol{x}}\|_G{}^2+|\lambda|^2\|\boldsymbol{z}\|_G{}^2+2\,\mathrm{Re}[\lambda(\!(\boldsymbol{y}-\hat{\boldsymbol{x}},\boldsymbol{z})\!)_G]\\&\geqq\|\boldsymbol{y}-\hat{\boldsymbol{x}}\|_G{}^2\end{aligned} \tag{3.7.19}$$

でなければならない．$\lambda=\mu\mathrm{e}^{\mathrm{i}\sigma}$ (μ,σ：実数)とおいて(3.7.19)が任意の μ,σ に対して成り立つためには，

$$\mu^2\|\boldsymbol{z}\|_G{}^2+2\mu\,\mathrm{Re}[\mathrm{e}^{\mathrm{i}\sigma}(\!(\boldsymbol{y}-\hat{\boldsymbol{x}},\boldsymbol{z})\!)_G]\geqq 0 \tag{3.7.20}$$

が任意の μ について成り立たねばならないから，

$$\mathrm{Re}[\mathrm{e}^{\mathrm{i}\sigma}(\!(\boldsymbol{y}-\hat{\boldsymbol{x}},\boldsymbol{z})\!)_G]=0. \tag{3.7.21}$$

(3.7.21)が任意の σ について成り立つためには

$$(\!(\boldsymbol{y}-\hat{\boldsymbol{x}},\boldsymbol{z})\!)_G=0, \tag{3.7.22}$$

すなわち，$\boldsymbol{y}-\hat{\boldsymbol{x}}$ は $\forall\boldsymbol{z}\in U$ と直交しなければならない．ゆえに $\boldsymbol{y}-\hat{\boldsymbol{x}}\in U^{\perp}$．このように，$\boldsymbol{y}=\hat{\boldsymbol{x}}+(\boldsymbol{y}-\hat{\boldsymbol{x}})$ は，$\hat{\boldsymbol{x}}\in U$，$\boldsymbol{y}-\hat{\boldsymbol{x}}\in U^{\perp}$ であるから，直和分解 $V_n=U\oplus U^{\perp}$ に対応しており，$\hat{\boldsymbol{x}}=P_{U,U^{\perp}}\boldsymbol{y}$ であることがわかる．

演 習 問 題

3.1 次の集合が，それぞれ，適当に体 K を選び，ベクトル和とスカラー倍を適当に定義すれば，ベクトル空間をなすことを確かめよ．また，空間の次元を定め，有限次元のものについては，具体的に基底を(少なくとも一つ，できれば有意義なものを二つ以上)挙げよ．

(i) 普通の“空間”の中の有向直線分の全体に対して，平行移動により一致するものを同一視したときの同値類の集合(“幾何ベクトル”と呼ぶことがある)．

(ii) 区間$[-1,1]$の上の複素数値連続関数の全体．

(iii) 体 K の元を係数とする 1 変数の n 次以下の多項式の全体 $K_n[x]$．

(iv) 体 K の上の $1\times n$ 型行列(横ベクトル)の全体，あるいは $m\times 1$ 型行列(縦ベクトル)の全体(“数ベクトル”と呼ぶことがある)．

(v)　体 K の上の $m\times n$ 型の行列の全体 $\mathscr{M}_{m,n}$.

(vi)　実数の無限列 $(x_0, x_1, x_2, \cdots)$ $(x_i\in\mathbb{R})$ の全体.

(vii)　与えられた $k+1$ 個の実数 $a_0, a_1, \cdots, a_k$ $(a_0\neq 0)$ に対して，"差分方程式" $a_0x_{n+k}+a_1x_{n+k-1}+\cdots+a_kx_n=0$ $(n=0,1,2,\cdots)$ を満たす実数列 $(x_0, x_1, x_2, \cdots)$ の全体.

(viii)　与えられた接続(位相構造)をもつ閉じた電気回路網の中の Kirchhoff(キルヒホッフ)の電流法則(だけ)を満たす電流配置として可能なものの全体. 同じく Kirchhoff の電圧法則(だけ)を満たす電圧配置として可能なものの全体.

(ix)　与えられた構造の平面トラスに一定の外力が作用しているとき，力の釣合い条件(だけ)を満たす部材力として可能なものの全体.

[略解]　ベクトルの和，スカラー倍は常識的に定義すればよい. それぞれの空間の次元は，(i)3，(ii)無限，(iii)$n+1$，(iv)n あるいは m，(v)$m\cdot n$，(vi)無限，(vii)k，(viii)回路の接続構造を表すグラフの零度(＝辺の数－階数)，および階数(＝頂点の数－連結成分の数)，(ix)トラスの不静定度.

3.2　F がマトロイド(→注意 3.1.4)のとき，F の任意の有限部分集合 E もマトロイドになる. E に関し，次が成り立つことを示せ.

(i)　$X(\subseteqq E)$ の極大独立部分集合の大きさを $\rho(X)$ と記し，X の**階数**(rank)と呼ぶ. このとき，E の有限部分集合に非負整数を対応づける階数 ρ は以下の⟨ρ0⟩,⟨ρ1⟩,⟨ρ2⟩を満たす.

⟨ρ0⟩　任意の $X(\subseteqq E)$ に対して $0\leqq\rho(X)\leqq|X|$.

⟨ρ1⟩　$X\subseteqq Y(\subseteqq E)$ ならば $\rho(X)\leqq\rho(Y)$.

⟨ρ2⟩　任意の $X, Y(\subseteqq E)$ に対して $\rho(X)+\rho(Y)\geqq\rho(X\cup Y)+\rho(X\cap Y)$.

(ii)　E の極大独立部分集合を E の**基**(base)と呼び，基の集合を $\mathscr{B}$ と記すと，$\mathscr{B}$ は以下の⟨B0⟩,⟨B1⟩を満たす.

⟨B0⟩　$\mathscr{B}\neq\emptyset$.

⟨B1⟩　$B_1, B_2\in\mathscr{B}$，$\boldsymbol{x}_1\in B_1-B_2$ ならば，ある $\boldsymbol{x}_2$ $(\in B_2-B_1)$ が存在して，

$$(B_1-\{\boldsymbol{x}_1\})\cup\{\boldsymbol{x}_2\}\in\mathscr{B}.$$

(iii)　$X(\subseteqq E)$ に対してその**閉包**(closure) $\mathrm{cl}(X)$ を，上記(i)における階数関数 ρ を用いて $\mathrm{cl}(X)=\{\boldsymbol{x}\in E\mid\rho(X)=\rho(X\cup\{\boldsymbol{x}\})\}$ により定義する. このとき cl は以下の⟨cl 0⟩,⟨cl 1⟩,⟨cl 2⟩を満たす.

⟨cl 0⟩　任意の $X(\subseteqq E)$ に対して，$X\subseteqq\mathrm{cl}(X)$.

⟨cl 1⟩　任意の $X, Y(\subseteqq E)$ に対して $X\subseteqq\mathrm{cl}(Y)$ ならば $\mathrm{cl}(X)\subseteqq\mathrm{cl}(Y)$.

⟨cl 2⟩　任意の $X(\subseteqq E)$ と $\boldsymbol{x}(\in E)$ に対して，

$$\boldsymbol{y}\in\mathrm{cl}(X\cup\{\boldsymbol{x}\})-\mathrm{cl}(X) \text{ ならば } \boldsymbol{x}\in\mathrm{cl}(X\cup\{\boldsymbol{y}\})-\mathrm{cl}(X).$$

(iv)　E の極小従属部分集合を**サーキット**(circuit)と呼び，サーキットの集合を $\mathscr{C}$ と記す(E の部分集合がみな独立であれば $\mathscr{C}=\emptyset$ である). このとき $\mathscr{C}$ は以下の⟨C0⟩,

$\langle C1\rangle$, $\langle C2\rangle$を満たす．

$\langle C0\rangle$　$\varnothing \notin \mathscr{C}$.

$\langle C1\rangle$　$C_1, C_2 \in \mathscr{C}$，$C_1 \subseteqq C_2$ ならば $C_1 = C_2$.

$\langle C2\rangle$　$C_1, C_2 \in \mathscr{C}$，$C_1 \neq C_2$，$\boldsymbol{x} \in C_1 \cap C_2$ ならば，$C_3 \subseteqq (C_1 \cup C_2) - \{\boldsymbol{x}\}$であるような $C_3(\in \mathscr{C})$が存在する．

逆に，有限集合 E の部分集合に非負整数を対応させる関数 ρ が(i)の$\langle \rho 0\rangle$, $\langle \rho 1\rangle$, $\langle \rho 2\rangle$を満たすとき，$\rho(X) = |X|$ なる $X(\subseteqq E)$を“独立”，$\rho(X) < |X|$ なる X を“従属”と呼べば E は一つのマトロイドとなることを示せ．さらに，マトロイド E の階数関数 ρ からこのようにして再定義されるマトロイドはもとのマトロイドに一致することも示せ．(ii)，(iii)，(iv)についても同様のことを示せ．

［略解］ この演習問題は“難しい”わけではないが，定義に基づいて「こつこつと」作業して行かなければならないので大変である．マトロイドのことを体系的に述べた本などを参照されたい．なお，(ii)の場合は「$\exists B \in \mathscr{B}: X \subseteqq B$」なる X を“独立”と呼ぶ．(iii)の場合は，「$\forall \boldsymbol{x} \in X: \mathrm{cl}(X - \{\boldsymbol{x}\}) \neq \mathrm{cl}(X)$」なる X を“独立”と呼ぶ．(iv)の場合は，「$\forall Y \subseteqq X: Y \notin \mathscr{C}$」なる X を“独立”と呼ぶ．

3.3　n 次元ベクトル空間 V_n を，二つのベクトル $\boldsymbol{x}, \boldsymbol{y} \in V_n$ の間の距離が $\|\boldsymbol{x} - \boldsymbol{y}\|_\infty$ で定義される“距離空間”であるとみなすとき，任意のノルム $\|\cdot\|: V_n \to \mathbb{R}_+$ は連続関数であることを示せ．

［略解］ V_n に基底 $B = \{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$を定め，$M = \sum_{i=1}^{n} \|\boldsymbol{e}_i\|$ とおく．任意の $\boldsymbol{x}(\in V_n)$，$\varepsilon(>0)$に対して，$\boldsymbol{x}$ の ε/M 近傍 $U(\boldsymbol{x}, \varepsilon/M) = \{\boldsymbol{z} \in V_n \mid \|\boldsymbol{z} - \boldsymbol{x}\|_\infty < \varepsilon/M\}$の中に $\boldsymbol{y}$ を取ると，3.1.5 項のノルムの性質(2)，(3)により，$|\,\|\boldsymbol{y}\| - \|\boldsymbol{x}\|\,| \leqq \|\boldsymbol{y} - \boldsymbol{x}\| \leqq \sum_{i=1}^{n} |y_i - x_i| \cdot \|\boldsymbol{e}_i\| \leqq \left(\max_i |y_i - x_i|\right) \cdot \sum_i \|\boldsymbol{e}_i\| = \|\boldsymbol{y} - \boldsymbol{x}\|_\infty \cdot M < \varepsilon$.

3.4　n 次元空間 V_n の任意の二つのノルム $\|\cdot\|$，$\|\cdot\|'$ は等価であることを示せ．

［略解］ 任意のノルムが最大値ノルム $\|\cdot\|_\infty$ に等価であることを示せば十分である．上の演習問題 3.3 のように V_n に位相(距離)を導入すると，n 次元箱 $S = \{\boldsymbol{x} \in V_n \mid \|\boldsymbol{x}\|_\infty = 1\}$は有界閉集合である．$\|\cdot\|$ は連続関数であるから，S の上での最大値 $b = \max_{x \in S} \|\boldsymbol{x}\|$，最小値 $a = \min_{x \in S} \|\boldsymbol{x}\|$ が存在する．ノルムの性質(1)により $a \neq 0, b \neq 0$．任意の非零ベクトル $\boldsymbol{y}$ に対して $\boldsymbol{y}/\|\boldsymbol{y}\|_\infty \in S$ であるから $a \leqq \|(\boldsymbol{y}/\|\boldsymbol{y}\|_\infty)\| \leqq b$．ゆえに $a\|\boldsymbol{y}\|_\infty \leqq \|\boldsymbol{y}\| \leqq b\|\boldsymbol{y}\|_\infty$.

3.5　ベクトル空間 V とその部分空間 U が与えられたとき，V の双対空間 V^{d} において $W^{\mathrm{d}} = \{\boldsymbol{y} \in V^{\mathrm{d}} \mid \forall \boldsymbol{x} \in U: \langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0\}$ は部分ベクトル空間をなすことを示せ．V における U の補空間の一つ W に対して同様にして定義される V^{d} の部分空間 $U^{\mathrm{d}} = \{\boldsymbol{y} \in V^{\mathrm{d}} \mid \forall \boldsymbol{x} \in W: \langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0\}$が V^{d} における W^{d} の補空間をなすことも示せ．さらに，$\{\boldsymbol{x} \in V \mid \forall \boldsymbol{y} \in W^{\mathrm{d}}: \langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0\}$が U に一致することを示せ．

［略解］ 定義から直ちに導かれることではあるが，基底を具体的に定めて調べるとよりよく諸関係が見透かせるかもしれない．

3.6　階数 r の n 次非負定値 Hermite 行列 A は，ある $r\times n$ 型の行列 B($\mathrm{rank}\ B=r$)を用いて $A=B^{\mathrm{H}}B$ と表せることを示せ(逆は明らか)．A が正定値の場合には，B は左からユニタリ行列を掛ける自由度を除いて一意に定まることを示せ．

［略解］　前半は，2.15.4 項 a，b より明らか；すなわち，D が非負対角項からなる対角行列であるから $D=\sqrt{D}^{\mathrm{H}}\sqrt{D}$ のように書ける．そこで $B=\sqrt{D}\,L$ とおけばよい．後半については，$D=LAL^{\mathrm{H}}=LB^{\mathrm{H}}BL^{\mathrm{H}}$($D=\mathrm{diag}[d_1,\cdots,d_n],d_i>0$)なので，$\sqrt{D}=\mathrm{diag}[\sqrt{d_1},\cdots,\sqrt{d_n}]$とおけば，$\sqrt{D}^{-1}LB^{\mathrm{H}}BL^{\mathrm{H}}\sqrt{D}^{-1}=I_n$ であるが，これは $Q=BL^{\mathrm{H}}\sqrt{D}^{-1}$ とおけば Q がユニタリであることを表している．すなわち，このような任意の B は $B=Q\sqrt{D}\,L^{\mathrm{H}-1}$ と表せる．

3.7　演習問題 3.6 の後半において，B を非負定値 Hermite 行列に選ぶことができることを示せ(このとき $A=B^2$ となる)．

［略解］　5.2 節の結果を援用し，$A=QDQ^{\mathrm{H}}$ (Q：ユニタリ行列，$D=\mathrm{diag}[\lambda_1,\cdots,\lambda_r,0,\cdots,0]$)($\lambda_i>0, r=\mathrm{rank}\ A$)と表す．$B=Q\sqrt{D}\,Q^{\mathrm{H}}$ は明らかに非負定値 Hermite 行列で，$B^2=A$．

3.8　n 次 Hermite 行列 A が正定値ならば，A を $\begin{array}{|c|c|}\hline A_{11} & A_{12}\\\hline A_{21} & A_{22}\\\hline\end{array}$ (A_{11}：$m\times m$，A_{22}：$(n-m)\times(n-m)$，A_{12}：$m\times(n-m)$，A_{21}：$(n-m)\times m$)とブロックに分けたとき，$A_{22}-A_{21}A_{11}{}^{-1}A_{12}$ も正定値であることを示せ．

［略解］　具体的な合同変換をおこなってみてもよいし，注意 3.5.13 と 2.10.2 項(vi)(式(2.10.9))とを組み合わせてもよい．

3.9　正定値 Hermite 行列の Cholesky 分解(→2.15.4 項)においては，行や列の順序入れ換えは不必要であることを示せ．

［略解］　注意 3.5.13 と(2.15.34)による．

3.10　n 次正定値 Hermite 行列 $A=[a_{ij}]$においては，$a_{ii}>0$ ($i=1,\cdots,n$)であり，Cholesky 分解の途中に現れるすべての行列 $A^{(k)}$ において，$\max\limits_{i=k,\cdots,n} a_{ii}^{(k)}=\max\limits_{i,j=k,\cdots,n}|a_{ij}^{(k)}|$ であることを示せ．［注：このことは，正定値行列がある意味で“対角優位”であることを述べている．］

［後半の略証］　もし $|a_{i_0j_0}^{(k)}|>\max\limits_{i} a_{ii}^{(k)}$ なら $|a_{i_0j_0}^{(k)}|>a_{i_0i_0}^{(k)},a_{j_0j_0}^{(k)}$ であり，主小行列式 $|A^{(k)}|_{i_0j_0}^{i_0j_0}=a_{i_0i_0}^{(k)}\cdot a_{j_0j_0}^{(k)}-a_{i_0j_0}^{(k)}\cdot a_{j_0i_0}^{(k)}=a_{i_0i_0}^{(k)}\cdot a_{j_0j_0}^{(k)}-|a_{i_0j_0}^{(k)}|^2<0$ となり，$A^{(k)}$ の正定値性に反する．

3.11　n 次正定値 Hermite 行列の Cholesky 分解の途中に現れる行列 $A^{(k)}$ に対しては，$(0<)\,a_{ii}^{(k)}\leqq a_{ii}^{(k-1)}$，$\max\limits_{i,j=k,\cdots,n}|a_{ij}^{(k)}|\leqq\max\limits_{i,j=k-1,\cdots,n}|a_{ij}^{(k-1)}|$ が成り立つ．

［略解］　正定値行列の LU 分解，Cholesky 分解は，行列要素の絶対値が常にある限度内に抑えられているという意味で“安定”であることを述べている．後半は前半と演習問題 3.10 の後半より．前半は $a_{ii}^{(k)}=a_{ii}^{(k-1)}-a_{ik}^{(k-1)}a_{ki}^{(k-1)}/a_{kk}^{(k-1)}=a_{ii}^{(k-1)}-|a_{ik}^{(k-1)}|^2/a_{kk}^{(k-1)}$

$\leqq a_{ii}^{(k-1)}$ ($\because a_{kk}{}^{(k-1)}>0$) による．

3.12　$Y=\{\boldsymbol{y}_1, \cdots, \boldsymbol{y}_m\}$が線形独立であるとき，

$$\boldsymbol{x}_s = s!\langle\!\langle \boldsymbol{y}_{[1}, \boldsymbol{y}_{|1|}\rangle\!\rangle_G \cdot \langle\!\langle \boldsymbol{y}_2, \boldsymbol{y}_{|2|}\rangle\!\rangle_G \cdot \cdots \cdot \langle\!\langle \boldsymbol{y}_{s-1}, \boldsymbol{y}_{|s-1|}\rangle\!\rangle_G \cdot \boldsymbol{y}_{s]}$$

により定義される Y の線形結合のベクトルの集合 $X=\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m\}$は直交系をなすことを示せ．[注：これは 3.6 節の Gram-Schmidt の直交化法と同じ内容であるが，より陽な表現であるともみなせる．]

[略解]

$$\langle\!\langle \boldsymbol{x}_s, \boldsymbol{y}_j\rangle\!\rangle_G = \det\begin{bmatrix} \langle\!\langle \boldsymbol{y}_1, \boldsymbol{y}_1\rangle\!\rangle_G & \cdots & \langle\!\langle \boldsymbol{y}_1, \boldsymbol{y}_{s-1}\rangle\!\rangle_G & \langle\!\langle \boldsymbol{y}_1, \boldsymbol{y}_j\rangle\!\rangle_G \\ \vdots & & \vdots & \vdots \\ \langle\!\langle \boldsymbol{y}_s, \boldsymbol{y}_1\rangle\!\rangle_G & \cdots & \langle\!\langle \boldsymbol{y}_s, \boldsymbol{y}_{s-1}\rangle\!\rangle_G & \langle\!\langle \boldsymbol{y}_s, \boldsymbol{y}_j\rangle\!\rangle_G \end{bmatrix}$$

であるから，$\langle\!\langle \boldsymbol{x}_s, \boldsymbol{y}_j\rangle\!\rangle_G=0$ $(j<s)$，$\langle\!\langle \boldsymbol{x}_s, \boldsymbol{x}_s\rangle\!\rangle_G>0$ (→2.16.2 項，Gram 行列式)．$L(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_j) \subseteqq L(\boldsymbol{y}_1, \cdots, \boldsymbol{y}_j)$ (実際には＝) であるから，$\langle\!\langle \boldsymbol{x}_s, \boldsymbol{x}_j\rangle\!\rangle_G=0$ $(j<s)$．

3.13　計量空間 V_n の部分空間 U_r に対して，その直交補空間 $U_r{}^{\perp}$ は一意に定まることを示せ．

[略解]　注意 3.6.2 のようにして構成した補空間を W とする．$\boldsymbol{y} \in V_n$ が $\forall \boldsymbol{x} \in U_r$ に直交するための必要十分条件は，$\boldsymbol{y}$ が U_r のすべての基底ベクトル $\boldsymbol{e}_i$ と直交すること：$\langle\!\langle \boldsymbol{y}, \boldsymbol{e}_i\rangle\!\rangle_G=0$ である．$\boldsymbol{y}=\sum_{i=1}^{r} y_i\boldsymbol{e}_i + \sum_{j=1}^{n-r} \tilde{y}_{r+j}(\boldsymbol{e}_{r+j})_{\perp}$ と表せば，これは $y_i=0$ $(i=1, \cdots, r)$ に他ならない．すなわち $\boldsymbol{y} \in W$．したがって $W=\{\boldsymbol{y} \in V_n \mid \forall \boldsymbol{x} \in U_r : \langle\!\langle \boldsymbol{y}, \boldsymbol{x}\rangle\!\rangle_G=0\}$．

3.14　計量空間 V_n において，部分空間 U の直交補空間を $U^{\perp}$ と記すと，次の関係が成り立つことを示せ．

(i)　$(U^{\perp})^{\perp}=U$，　(ii)　$(U_1+U_2)^{\perp}=U_1{}^{\perp} \cap U_2{}^{\perp}$，　(iii)　$(U_1 \cap U_2)^{\perp}=U_1{}^{\perp}+U_2{}^{\perp}$．

[略解]　直交補空間の定義から直ちに導かれる．演習問題 3.5 も参照．

4
線形方程式系

線形方程式系(system of linear equations, linear system of equations) は古くは連立一次方程式(simultaneous linear equations, linear simultaneous equations)とも呼ばれ，歴史的にも応用上も，線形代数の中で最も人口に膾炙している形態の主題である．

線形方程式系の理論展開に必要な諸概念も解法技術の基本的なものも，実質的にはすでに前章までに論じ尽くされているので，本章ではそれらを線形方程式系の言葉でまとめて述べ直すに止める．その他，応用上も重要と思われる二三の細かな事項をつけ加える．

一々注意はしないが，本章の話題は「線形計算」と呼ばれる数値計算の分野と非常に関係が深い．数値計算により方程式を解く際の細かな技巧や注意事項についてはその分野の専門書を参照されたい．

4.1 線形方程式系

$m\times n$ 型の行列 $A=[a_{ij}]$ と m 次元縦ベクトル $\boldsymbol{b}=[b_i]$ が与えられたとき

$$A\boldsymbol{x}=\boldsymbol{b}: \quad \sum_{j=1}^{n} a_{ij}x_j=b_i \qquad (i=1,\cdots,m) \tag{4.1.1}$$

を満たす未知ベクトル(n 次元縦ベクトル) $\boldsymbol{x}=[x_j]$ を求める問題を本章で扱う．

(4.1.1)のことを，A を**係数行列**(coefficient matrix)，$\boldsymbol{b}$ を**右辺**(**ベクトル**)(right-hand side)，$\boldsymbol{x}$ を**未知数**(**未知ベクトル**；unknown(s))とする(n 次元の)**線形方程式系**((m 元の)**連立一次方程式**)と呼ぶ．

もちろん，A が n 次正則行列であれば，(4.1.1)には一意解 $\boldsymbol{x}=A^{-1}\boldsymbol{b}$ が存在する．

4.2　有　理　解　法

線形方程式系の解を，所与量(係数行列の mn 個の要素と右辺ベクトルの m 個の要素)に対して有限回の四則演算(ときには開平も許すことがある)を施すことによって求める方法のことを“有理解法”と呼ぶことにする．

注意 4.2.1　有理解法は 4.4 節で述べる反復法と対置させられるもので，**直接(解)法**と呼ばれることも多い．

線形方程式系の解の公式として歴史的に有名であるのは **Cramer**(クラメール)**の公式**(Cramer's rule)である．それは A が n 次正則のとき(4.1.1)の解 $\boldsymbol{x}=A^{-1}\boldsymbol{b}$ が

$$x_j=\det\begin{bmatrix} a_{11} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{nn} \end{bmatrix} \Big/ \det A \quad (j=1,\cdots,n) \tag{4.2.1}$$

で表されるというものである．これが正しいことを示すには，A が正則($\det A \neq 0$)のときには A^{-1} が存在し，$\boldsymbol{x}=A^{-1}\boldsymbol{b}$ が一意に確定するから，(4.2.1)の $\boldsymbol{x}$ が(4.1.1)を満たすことを示せばよい．これは，2.8 節の(列に関する)Laplace 展開(2.8.2)を(4.2.1)の右辺の分子の行列式に適用した式(A の余因子行列 $\hat{A}=\mathrm{adj}\,A=[\hat{a}_{ji}]$ とする；(2.10.1)も参照)

$$x_j=\sum_{i=1}^{n}\hat{a}_{ji}b_i/\det A, \quad \text{すなわち}\ \boldsymbol{x}=\hat{A}\boldsymbol{b}/\det A=A^{-1}\boldsymbol{b} \tag{4.2.2}$$

より明らかである．

注意 4.2.2　Cramer の公式は有名であり，確かに応用上も問題の性質を理論的に論じる際に役に立つ場合もあるが，$n=3$ ぐらいまでならともかくそれより変数の数 n の大きい方程式を“数値的”に解くのには適さないということはいまでは常識となっている．

2.15 節で述べた行列の各種の基本的な変形は，線形方程式系を解くための実用的な方法も与えてくれる．

2.15.1 項“列の置換と行に関する基本変形”によって，係数行列 A を(2.15.1)の形 $\tilde{A}=SAP$ に変形し，同時に右辺ベクトル $\boldsymbol{b}$ に対しても同じ基本行演算を施し $\tilde{\boldsymbol{b}}=S\boldsymbol{b}$ を作れば，(4.1.1)は

$$\tilde{A}P^{-1}\boldsymbol{x}=SAP(P^{-1}\boldsymbol{x})=S\boldsymbol{b}=\tilde{\boldsymbol{b}} \tag{4.2.3}$$

となるから(P は A の列の置換を表す置換行列)，$\tilde{\boldsymbol{x}}=P^{-1}\boldsymbol{x}$ ($\tilde{\boldsymbol{x}}$ の座標成分に A の列の置換と同じ置換を施したものが $\boldsymbol{x}$ となる)とおくと

$$\left[\begin{array}{cc|c} 1 & O & \\ & \ddots & C \\ O & 1 & \\ \hline & O & \end{array}\right] \begin{bmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_r \\ \tilde{x}_{r+1} \\ \vdots \\ \tilde{x}_n \end{bmatrix} = \begin{bmatrix} \tilde{b}_1 \\ \vdots \\ \tilde{b}_m \end{bmatrix} \tag{4.2.4}$$

という形に表される($r=\mathrm{rank}\,A$)．このことから，(4.1.1)は(4.2.4)において

$$\tilde{b}_i=0 \qquad (i=r+1,\cdots,m) \tag{4.2.5}$$

のとき，そしてそのときに限り解をもち，その解は，$\tilde{x}_{r+1},\cdots,\tilde{x}_n$ を任意パラメタとして

$$\tilde{x}_i=\tilde{b}_i-\sum_{j=1}^{n-r}c_{ij}\tilde{x}_{r+j} \qquad (i=1,\cdots,r) \tag{4.2.6}$$

という形に表現できる(c_{ij} は(4.2.4)の左辺の係数行列の中の部分行列 C の要素)．

注意 4.2.3 (4.2.5)の条件が成立していないときには方程式に解がない．これは，“条件が厳しすぎる”ことを表しているともみなせるので，この場合，方程式系は**条件過剰**(overspecified, overdetermined)であるという(古くは**不能**とも呼ばれていた)．

(4.2.6)の右辺に $n-r$ 個の任意パラメタ $\tilde{x}_{r+j}$ が残るとき($n-r>0$ のとき)には，解が一意に確定しないという意味で，方程式系は**条件不足**(underspecified, underdetermined)であるという(古くはこの場合は**不定**と呼ばれていた)．

なお，注意 3.3.4 も参照のこと．

注意 4.2.4 $n \doteqdot m \doteqdot r$ のとき，このようなやり方で解を計算するのに要する四則演算は，ほぼ $n^3/2$ 回の乗算と同程度の回数の加減算とそれに約 n 回の除算とである(→注意 2.15.5)．

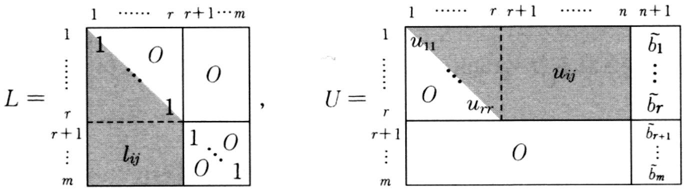
注意 4.2.5　この解法は線形計算の分野で **Gauss-Jordan**（ガウス・ヨルダン）[1] **法**と呼ばれているものと実質的には同じものである．

方程式系(4.1.1)を，$\tilde{A}=[A, \boldsymbol{b}]$（$m\times(n+1)$型行列）とおくことによって

$$\tilde{A}\begin{bmatrix} x_1 \\ \vdots \\ x_n \\ -1 \end{bmatrix}=\mathbf{0} \tag{4.2.7}$$

と書き表し，$\tilde{A}$ に 2.15.3 項の LU 分解を施してみる．このとき $\tilde{A}$ の列 $n+1$（$\boldsymbol{b}$ の列）は他の列と置換を受けないようにする．すると，$\tilde{A}$ に行の置換 P_1，列の置換 P_2 を施したものの LU 分解(2.15.48)が

$$P_1\tilde{A}P_2=LU,$$



(4.2.8)

$(u_{11},\cdots,u_{rr}\neq 0)$ という形で得られる．そこで，

$$\begin{bmatrix} \tilde{\boldsymbol{x}} \\ -1 \end{bmatrix}=P_2^{-1}\begin{bmatrix} \boldsymbol{x} \\ -1 \end{bmatrix} \quad \text{あるいは} \quad \begin{bmatrix} \boldsymbol{x} \\ -1 \end{bmatrix}=P_2\begin{bmatrix} \tilde{\boldsymbol{x}} \\ -1 \end{bmatrix} \tag{4.2.9}$$

により $\tilde{\boldsymbol{x}}$ を定義すれば，(4.1.1)は

$$LU\begin{bmatrix} \tilde{\boldsymbol{x}} \\ -1 \end{bmatrix}=\mathbf{0} \tag{4.2.10}$$

という形に表される．$\det L=1\ (\neq 0)$ であるから，(4.2.10)は

$$U\begin{bmatrix} \tilde{\boldsymbol{x}} \\ -1 \end{bmatrix}=\mathbf{0} \tag{4.2.11}$$

と等価である．このことから，(4.1.1)は(4.2.8)において

$$\tilde{b}_i=0 \qquad (i=r+1,\cdots,n) \tag{4.2.12}$$

---

1)　この Jordan は Wilhelm Jordan（ヴィルヘルム・ヨルダン，1842～1899）で，解析学などで有名な Camille Jordan（カミーユ・ジョルダン，1838～1922）とは別人である．本書 6.8 節の Jordan 標準形に名を残しているのは，Camille Jordan の方である．また，本書では扱わないが，行列代数に関係の深い Jordan 代数の Jordan は，さらに別人の Pascual Jordan（パスクアル・ヨルダン，1902～1980）である．

のとき，そしてそのときに限り，解をもち，その解は，$x_j\,(j=r+1,\cdots,n)$を任意パラメタとして，$i:=r,r-1,\cdots,1$ の順に

$$\tilde{x}_i=\left(\tilde{b}_i-\sum_{j=i+1}^{n}u_{ij}\tilde{x}_j\right)\Big/u_{ii} \qquad (4.2.13)$$

を用いて $\tilde{x}_i$ $(i=1,\cdots,r)$を計算することによって定められる．

注意 4.2.6　$n\doteqdot m\doteqdot r$ のとき，LU 分解に要する四則演算は，ほぼ $n^3/3$ 回の乗算と加減算，それに約 n 回の除算である．(4.2.13)の計算には約 $n^2/2$ 回の乗算と加減算，それに約 n 回の除算(これは LU 分解のときと共用できる)が要る．

注意 4.2.7　この解法は，数値計算の分野では **Gauss の消去法**(Gaussian elimination)と呼ばれている．LU 分解の部分を**前進消去**(forward elimination)，(4.2.13)の計算の部分を**後退代入**(backward substitution)と呼んだりもする．

注意 4.2.8　係数行列 A が対称あるいは Hermite のときには，2.15.4 項の Cholesky(風)分解を利用して，上と同様の解法が作れるがここでは詳しくは述べない．計算の手間は LU 分解を利用するものの半分になる．

注意 4.2.9　2.15.5 項の QR 分解を利用する解法についても同様な議論ができる．

4.3　解の存在と一意性

すでに 4.2 節において，解法手続きの中で線形方程式系の解の存在や一意性については述べてしまった．以下では，ベクトル空間の言葉を使ってそれらを述べ直し，いくつかの注意を与える．

4.3.1　解の存在条件

線形方程式系(4.1.1)の係数行列 $A=[a_{ij}]=[\boldsymbol{a}_{\cdot 1},\boldsymbol{a}_{\cdot 2},\cdots,\boldsymbol{a}_{\cdot n}]$の n 本の列ベクトルを $\boldsymbol{a}_{\cdot j}\,(j=1,\cdots,n)$とすると，(4.1.1)は

$$x_1\boldsymbol{a}_{\cdot 1}+\cdots+x_n\boldsymbol{a}_{\cdot n}=\boldsymbol{b} \qquad (4.3.1)$$

と書けるから，別のいい方をすれば，(4.1.1)は「右辺ベクトル $\boldsymbol{b}$ を係数行列の列ベクトルの線形結合として表す」問題，すなわち

$$「\boldsymbol{b}\in L(\boldsymbol{a}_{\cdot 1},\cdots,\boldsymbol{a}_{\cdot n})\text{であるか？}」 \qquad (4.3.2)$$

とみなすことができる．あるいは，3.3.3 項の言葉を使って，線形写像 A：$V_n\to U_m$ を考えれば，

$$「\boldsymbol{b} \in \mathrm{Im}\, A\ \text{であるか？}」 \tag{4.3.3}$$

と述べることができる．さらに，A の列に $\boldsymbol{b}$ を添加した行列 $[A, \boldsymbol{b}]$ を考えれば，一般には $\mathrm{rank}\ A \leqq \mathrm{rank}[A, \boldsymbol{b}]$ であるが，(4.3.2)が成り立つかどうかは

$$\mathrm{rank}\ A = \mathrm{rank}[A, \boldsymbol{b}] \tag{4.3.4}$$

が成り立つかどうかともいい直せる．すなわち「線形方程式系(4.1.1)に解が存在するための必要十分条件は，(4.3.2)，(4.3.3)，(4.3.4)(いずれも等価な条件)が成立することである．」

U_m の双対空間 U_m^{d} の言葉を使って(4.3.2)，(4.3.3)の条件を述べると(演習問題3.5参照)，

「$\langle \boldsymbol{a}_{\cdot j}, \boldsymbol{y} \rangle = 0\ (j=1,\cdots,n)$ を満たす(同じことだが $\boldsymbol{y}^{\mathsf{T}} A = \boldsymbol{0}$ を満たす)すべての $\boldsymbol{y} \in U_m^{\mathrm{d}}$ に対して $\langle \boldsymbol{b}, \boldsymbol{y} \rangle = 0$」　(4.3.5)

となる．この形も有用なことがある．

注意 4.3.1　(4.3.5)と(4.3.2)が等価であることを述べる定理は **Farkas**(ファルカシュ)**型の定理**の一つともみなせる．転置を取った行列形式で述べると次の通り．

[Farkas 型の定理—線形等式版(K は任意の体)]

与えられた A と $\boldsymbol{c}$ に対して，条件

$$「\text{任意の}\ \boldsymbol{x}\ \text{に対して，}A\boldsymbol{x}=\boldsymbol{0}\ \text{ならば}\ \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x}=0」 \tag{4.3.6}$$

と条件

$$「\boldsymbol{c}^{\mathsf{T}} = \boldsymbol{y}^{\mathsf{T}} A\ \text{を満たす}\ \boldsymbol{y}\ \text{が存在する}」 \tag{4.3.7}$$

とは等価である．

Farkas 型の定理としてよく知られているものをさらに二つ挙げておく．(証明略；不等式版については「線形計画法」関係の専門書を参照．整数版については2.15.6項をもとにして証明できる；6.6.3項参照．)

[Farkas の定理—線形不等式版($K=\mathbb{R}$(実数体))]

与えられた A と $\boldsymbol{c}$ に対して，条件

$$「\text{任意の}\ \boldsymbol{x}\ \text{に対して，}A\boldsymbol{x} \geqq \boldsymbol{0}\ \text{ならば}\ \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x} \geqq 0」 \tag{4.3.8}$$

と条件

$$「\boldsymbol{c}^{\mathsf{T}} = \boldsymbol{y}^{\mathsf{T}} A\ \text{を満たす}\ \boldsymbol{y}(\geqq \boldsymbol{0})\ \text{が存在する}」 \tag{4.3.9}$$

とは等価である．ここでは，ベクトルの間の不等号“≧”は各成分ごとに不等式関係があることを意味するものとする．

[Farkas 型の定理—整数版($K=\mathbb{Q}$(有理数体))]

与えられた A と $\boldsymbol{c}$ に対して，条件

$$「\text{任意の}\ \boldsymbol{x}\ \text{に対して，“}A\boldsymbol{x}\ \text{の全成分が整数”ならば}\ \boldsymbol{c}^{\mathsf{T}}\boldsymbol{x}\ \text{が整数}」 \tag{4.3.10}$$

と条件

$$「\boldsymbol{c}^{\top}=\boldsymbol{y}^{\top}A \text{を満たす整数ベクトル} \boldsymbol{y} \text{が存在する」} \tag{4.3.11}$$

とは等価である．

4.3.2　解の一意性条件

(4.1.1)に二つの解 $\boldsymbol{x}, \boldsymbol{x}'$ があれば $\boldsymbol{y}=\boldsymbol{x}'-\boldsymbol{x}$ は同次方程式

$$A\boldsymbol{y}=\boldsymbol{0} \tag{4.3.12}$$

を満たす．したがって，(4.1.1)の解が(存在するときそれが)一意に定まるための必要十分条件は

$$A\boldsymbol{y}=\boldsymbol{0} \quad \text{ならば} \quad \boldsymbol{y}=\boldsymbol{0} \tag{4.3.13}$$

である．この条件は，線形写像の言葉(3.3.3項)を用いれば，

$$\mathrm{Ker}\ A=\{\boldsymbol{0}\} \tag{4.3.14}$$

あるいは

$$\mathrm{rank}\ A=n\,(=A \text{ の列の数}) \tag{4.3.15}$$

ともいい表される．

よく用いられる特殊な場合として「A が n 次正方行列のとき，$A\boldsymbol{y}=\boldsymbol{0}$ が非零解 $\boldsymbol{y}\,(\neq\boldsymbol{0})$ を有するための必要十分条件は $\det A=0$ である」という形がある．

例 4.3.1　$K=\mathbb{C}$ とする．二つの1変数多項式

$$\left.\begin{aligned} p(x)&=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0 \qquad (a_n\neq 0),\\ q(x)&=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_1x+b_0 \qquad (b_m\neq 0) \end{aligned}\right\} \tag{4.3.16}$$

が共通の零点 α をもてば，$p(\alpha)=q(\alpha)=0$ であるから，図4.3.1の $m+n$ 個の式が同時に0となる．

図4.3.1の関係は行列 R およびベクトル $\boldsymbol{\alpha}$ を，それぞれ，

$$\alpha^{m-1}p(\alpha)=a_n\alpha^{n+m-1}+a_{n-1}\alpha^{n+m-2}+\cdots\cdots+a_1\alpha^m+a_0\alpha^{m-1}$$

$$\alpha^{m-2}p(\alpha)=a_n\alpha^{n+m-2}+a_{n-1}\alpha^{n+m-3}+\cdots\cdots+a_1\alpha^{m-1}+a_0\alpha^{m-2}$$

$$\vdots$$

$$p(\alpha)=a_n\alpha^n+a_{n-1}\alpha^{n-1}+\cdots\cdots+a_1\alpha+a_0$$

$$\alpha^{n-1}q(\alpha)=b_m\alpha^{n+m-1}+b_{m-1}\alpha^{n+m-2}+\cdots\cdots+b_1\alpha^n+b_0\alpha^{n-1}$$

$$\alpha^{n-2}q(\alpha)=b_m\alpha^{n+m-2}+b_{m-1}\alpha^{n+m-3}+\cdots\cdots+b_1\alpha^{n-1}+b_0\alpha^{n-2}$$

$$\vdots$$

$$q(\alpha)=b_m\alpha^m+b_{m-1}\alpha^{m-1}+\cdots\cdots+b_1\alpha+b_0$$

図 4.3.1

$$
R=\begin{array}{r|cccccccccccc|}
 & 1 & 2 & \cdots & \cdots & \cdots & m+1 & \cdots & n+1 & \cdots & \cdots & \cdots & m+n \\
\hline
1 & a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_1 & a_0 & & & & \\
2 & & a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_1 & a_0 & & & \\
\vdots & & & \ddots & \ddots & & & & & \ddots & \ddots & & \\
\vdots & & & & \ddots & \ddots & & & & & \ddots & \ddots & \\
m & & & & & a_n & a_{n-1} & \cdots & \cdots & \cdots & \cdots & a_1 & a_0 \\
m+1 & b_m & b_{m-1} & \cdots & \cdots & b_1 & b_0 & & & & & & \\
m+2 & & b_m & b_{m-1} & \cdots & \cdots & b_1 & b_0 & & & & & \\
\vdots & & & \ddots & \ddots & & & \ddots & \ddots & & & & \\
\vdots & & & & \ddots & \ddots & & & \ddots & \ddots & & & \\
\vdots & & & & & \ddots & \ddots & & & \ddots & \ddots & & \\
\vdots & & & & & & \ddots & \ddots & & & \ddots & \ddots & \\
m+n & & & & & & & b_m & b_{m-1} & \cdots & \cdots & b_1 & b_0 \\
\hline
\end{array}
\tag{4.3.17}
$$

(列番号の m と n の相対位置は m と n の大小関係に依存する)および

$$\boldsymbol{\alpha}=[\alpha^{m+n-1},\alpha^{m+n-2},\cdots,\alpha,1]^{\top} \tag{4.3.18}$$

とおくと，

$$R\boldsymbol{\alpha}=\mathbf{0} \tag{4.3.19}$$

と表せるから，(4.3.19)を満たす $\boldsymbol{\alpha}\neq\mathbf{0}$ が存在することにより，$p(x)$ と $q(x)$ が共通零点をもつための“必要条件”として

$$\det R=0 \tag{4.3.20}$$

が得られる．

ところが，(4.3.20)は，$p(x)$ と $q(x)$ が共通零点をもつため“十分条件”でもある．なぜならば，$p(x)$ の零点を $\alpha_1,\cdots,\alpha_n$ とし，$q(x)$ の零点を $\beta_1,\cdots,\beta_m$ とすれば，

$$p(x)=a_n\prod_{i=1}^{n}(x-\alpha_i),\qquad q(x)=b_m\prod_{j=1}^{m}(x-\beta_j) \tag{4.3.21}$$

であるが，

$$\det R=a_n{}^m b_m{}^n\prod_{i=1}^{n}\prod_{j=1}^{m}(\alpha_i-\beta_j)=a_n{}^m\prod_{i=1}^{n}q(\alpha_i)=(-1)^{nm}b_m{}^n\prod_{j=1}^{m}p(\beta_j) \tag{4.3.22}$$

であるからである．((4.3.22)が成り立つことは，次のように考えれば理解され

よう．det R は $p(x)$, $q(x)$ の係数の多項式であるが，それらの係数は，Vieta (ウィエタ)[1] の関係式により a_n, b_m と $\alpha_i, \beta_j (i=1,\cdots,n; j=1,\cdots,m)$ の多項式である．したがって，det R を $a_n, b_m, \alpha_i, \beta_j$ の多項式とみなすことができる．そうしたとき，もしある i と j に対して $\alpha_i=\beta_j$ であれば det $R=0$ となるというのであるから，det R は $\prod_i\prod_j(\alpha_i-\beta_j)$ で割り切れる．det R も $\prod_i\prod_j(\alpha_i-\beta_j)$ も各 α_i に関して m 次，各 β_j に関して n 次であるから，α_i, β_j によらない c を用いて det $R=c\prod_i\prod_j(\alpha_i-\beta_j)$ と表せる．たとえば$\left(\prod_{j=1}^{m}(-\beta_j)\right)^n$ の係数を比較して，$c=a_n{}^m b_m{}^n$ であることがわかる．)

det R のことを，$p(x)$ と $q(x)$ の**終結式**(resultant)という．det $R=0$ は $p(x)=0$ と $q(x)=0$ とから“x を消去”した形であるとみなせるので，この過程を**消去過程**(elimination)とと呼ぶことがある．

なお，2 変数以上の多項式のいくつかが与えられたとき，それらの対を考え，各対に対して，一つの変数に着目して，他の変数を K に添加した多項式環を係数領域とする 1 変数多項式とみなし，det R を作ると，変数の数が 1 個少ない多項式が得られるので，“理論上は”次々と変数を“消去”して行くことができる．□

例 4.3.2　n 次多項式 $p(x)$ の係数と m 次多項式 $q(x)$ の係数から作られた行列 R(→例 4.3.1)の $m+n-2i$ 次の正方部分行列 $R_i=R_{i+1,\cdots\cdots\cdots\cdots,m+n-i}^{i+1,\cdots,m,m+i+1,\cdots,m+n}$ の行列式を($p(x)$ と $q(x)$ の) i 次の**部分終結式**(subresultant)という．さらに，R_i の最後の列を $[x^{m-i-1}p(x), x^{m-i-2}p(x), \cdots, p(x), x^{n-i-1}q(x), x^{n-i-2}q(x), \cdots, q(x)]^{\mathsf{T}}$ でおきかえて得られる行列 S_i の行列式も考える．S_i の第 1 列に $x^{m+n-i-1}$ を，第 2 列に $x^{m+n-i-2}$ を，…，第 $m+n-2i-1$ 列に x^{i+1} $(i+1=(m+n-i)-(m+n-2i-1))$ を，それぞれ掛けて，第 $m+n-2i$ 列(最右列)から引くと，最右列の成分は，たかだか i 次の x の多項式となる．したがって det S_i の x に関する次数もたかだか i である．一方，det S_i は，その定義から，$p(x)$ と $q(x)$ の最大公約多項式 $d(x)$ で割り切れる．そこで，$d(x)$ の次数が h ならば，det $S_i=0$ $(i=0,1,\cdots,h-1)$ で，det $S_h=c\cdot d(x)$ $(c\in K-\{0\})$ であるということになる．($c\neq 0$ であることは S_h の計算法から明らか．)ところで，$R_0=R$，det $S_0=$det R_0 であるから，$p(x)$ と $q(x)$ とが共通因子をもたなければ，$h=$

---

1)　フランスの数学者 Viète(ヴィエート)のラテン語名．この関係式はいわゆる“根(解)と係数の関係”である．

0 ($d(x)\in K$)で, $\det S_0=\det R_0=\det R\neq 0$ である. このようにして, 例 4.3.1 の結果は一般化される. (例 4.3.1 において(4.3.20)が, $p(x)$ と $q(x)$ が共通零点をもつための必要十分条件であることを示したが, ここでの議論はそのことの別証を与えていることにも注意されたい.) □

4.3.3 解空間の基底表現

すでに 4.2 節において, 線形方程式系(4.1.1) $A\boldsymbol{x}=\boldsymbol{b}$ (A は $m\times n$ 型) の "一般解" を $n-r$ 個($r=\mathrm{rank}\,A$)の未知数をパラメタとして表す式を得た(たとえば(4.2.6), (4.2.13)). このように, いくつかの自由に値を取りうる変数(座標成分)(その番号の集合を J とする)の 1 次式で残りの変数(その番号の集合を I とする)を表すことによって解空間のベクトルを表現する形式を**基底形式**(basic form)という($I\cup J=\{1,2,\cdots,n\}$である):

$$x_i+\sum_{j\in J}b_{ij}x_j=0 \qquad (i\in I). \tag{4.3.23}$$

一般には多くの基底形式があるが, その間の交換が 2.15.2 項に述べた**枢軸変換**である.

4.4 反 復 解 法

本節では $K=\mathbb{R}$ あるいは $K=\mathbb{C}$ とする.

4.4.1 線形反復解法とその収束性

一般にある方程式系

$$f_i(x_1,\cdots,x_n)=0 \qquad (i=1,\cdots,n) \tag{4.4.1}$$

の解 $\hat{\boldsymbol{x}}=[\hat{x}_1,\cdots,\hat{x}_n]^{\mathsf{T}}$ を求めるのに, ある**出発値**(**初期近似値**; starting value, initial approximation) $\boldsymbol{x}^{(0)}$ から出発し, **反復公式**(**漸化式**; iteration formula, recurrence relation)

$$\boldsymbol{x}^{(\nu+1)}=\varphi(\boldsymbol{x}^{(\nu)}) \qquad (\nu=0,1,2,\cdots) \tag{4.4.2}$$

を用いて, 解 $\hat{\boldsymbol{x}}$ に "収束" する系列 $\boldsymbol{x}^{(0)},\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\cdots$ を作り出すことを**反復法**(iterative method)という. ここで, φ は(一般には非線形の)写像 $V_n\to V_n$ であり, "収束" は適当なノルム(→3.1.5 項)に関して $\|\boldsymbol{x}^{(\nu)}-\hat{\boldsymbol{x}}\|\to 0$ ($\nu\to\infty$) を意

味することとする(どんなノルムを用いてもよい(→演習問題 3.4))．とくに $\varphi(\boldsymbol{x})=F\boldsymbol{x}+\boldsymbol{c}$ (F は n 次正方行列)の形のときに**線形反復**(linear iteration)という．

本節では線形方程式系(4.1.1)で $m=n$ の場合に対する線形反復解法について概観するが，実際に数値計算で解を求める際に注意すべきことがらや，手の込んだ具体的な公式などについては「線形計算」の専門書に譲り，ここでは省略する．また，線形反復の一般論には固有値(→第 5 章)や Jordan 標準形(→6.8 節)の話が必要なので，そこで再び述べることにする．

線形方程式系の反復解法の多くは次のような形式をとる．第 ν 近似(the ν-th approximation) $\boldsymbol{x}^{(\nu)}$ を $\boldsymbol{x}$ に代入したときの方程式系の左右両辺の喰い違いすなわち**残差**(residual)

$$\boldsymbol{r}^{(\nu)}=\boldsymbol{b}-A\boldsymbol{x}^{(\nu)} \tag{4.4.3}$$

を計算し，この残差にある線形変換 B(これが“反復公式”を特徴づける)を施したもの

$$\boldsymbol{d}^{(\nu)}=B\boldsymbol{r}^{(\nu)} \tag{4.4.4}$$

を**修正量**(correction)として，次の近似(第 $\nu+1$ 近似)

$$\boldsymbol{x}^{(\nu+1)}=\boldsymbol{x}^{(\nu)}+\boldsymbol{d}^{(\nu)} \tag{4.4.5}$$

を計算する．

注意 4.4.1　上記の反復の一段階は

$$\boldsymbol{x}^{(\nu+1)}=\boldsymbol{x}^{(\nu)}+B(\boldsymbol{b}-A\boldsymbol{x}^{(\nu)})=B\boldsymbol{b}+(I-BA)\boldsymbol{x}^{(\nu)} \tag{4.4.6}$$

という式にまとめることもできるが，実際の計算には(4.4.6)(の最右辺)を用いるのではなく，(4.4.3)～(4.4.5)を用いる．このようにすると，$\boldsymbol{r}^{(\nu)}$ や $\boldsymbol{d}^{(\nu)}$ の“大きさ”を観察することによって，近似度の改良され方が見えるし，残差の計算(4.4.3)だけを十分精密におこなえば全体として高精度の計算ができるからである．

いま，第 ν 近似 $\boldsymbol{x}^{(\nu)}$ と真の解 $\hat{\boldsymbol{x}}$ との差を

$$\boldsymbol{\varepsilon}^{(\nu)}=\hat{\boldsymbol{x}}-\boldsymbol{x}^{(\nu)} \tag{4.4.7}$$

と記せば，明らかに，

$$A\boldsymbol{\varepsilon}^{(\nu)}=\boldsymbol{r}^{(\nu)} \tag{4.4.8}$$

である．そして，(4.4.3)～(4.4.5)(あるいは(4.4.6))により

$$\boldsymbol{\varepsilon}^{(\nu+1)}=(I-BA)\,\boldsymbol{\varepsilon}^{(\nu)} \tag{4.4.9}$$

を得る．このことは，$I-BA$ が“小さい”ほど，すなわち B が A^{-1} に“近い”ほど，近似誤差 $\boldsymbol{\varepsilon}^{(\nu)}$ の減り方が速いことを示している．理論的には，6.8 節で述べるように，$I-BA$ の“固有値”の絶対値の最大値（$I-BA$ のスペクトル半径 $\rho(I-BA)$）が収束性，収束速度を支配するのであるが，大雑把に“ノルム”(→3.1.5 項，3.3.2 項)で表現すれば，(4.4.9)から

$$\|\boldsymbol{\varepsilon}^{(\nu+1)}\|\leqq\|I-BA\|\cdot\|\boldsymbol{\varepsilon}^{(\nu)}\| \tag{4.4.10}$$

を得るので，適当な(自然)ノルムに関して

$$\|I-BA\|<1 \tag{4.4.11}$$

ならば，$\nu\to\infty$ のとき $\|\boldsymbol{\varepsilon}^{(\nu)}\|\to 0$，すなわち $\boldsymbol{x}^{(\nu)}\to\hat{\boldsymbol{x}}$ であることになる(収束の“十分”条件)．

注意 4.4.2 反復解法が実用上重要であるのは，扱う問題が非常に大きいとき，その問題を“記述”するスペースさえあれば，それ以外にほとんど作業用のスペースを要しない(ように解法を設計することができる)ということにもある．直接法は，いろいろな工夫をする余地はあるにしても，問題を記述するスペースの他に，計算の作業用のスペースが，一般には，かなり必要となる．大規模線形計算においては，このような観点から解法を選択することが大切である．

4.4.2 Jacobi 反復

係数行列 A の対角要素は 0 でないとする：

$$a_{ii}\neq 0 \qquad (i=1,\cdots,n). \tag{4.4.12}$$

$\boldsymbol{x}^{(\nu)}$ から $\boldsymbol{x}^{(\nu+1)}$ を定めるときに，

「各 $i(=1,\cdots,n)$ ごとに，$j\neq i$ に対しては $x_j=x_j^{(\nu)}$ とおいて，$A\boldsymbol{x}=\boldsymbol{b}$ の第 i 式が成り立つように $x_i^{(\nu+1)}$ を独立に定める」

すなわち，

$$x_i^{(\nu+1)}=\frac{1}{a_{ii}}\left(b_i-\sum_{j\neq i}a_{ij}x_j^{(\nu)}\right) \qquad (i=1,\cdots,n) \tag{4.4.13}$$

とするのが **Jacobi 反復**(Jacobi iteration)と呼ばれるものである．

注意 4.4.3 方程式系 $A\boldsymbol{x}=\boldsymbol{b}$ の一つの方程式(i 番目の方程式)

$$\sum_{j=1}^{n}a_{ij}x_j=b_i \tag{4.4.14}$$

にだけ着目して他の方程式による束縛を“緩和(relax)”して，次の近似値を計算するというやり方を，すべての変数について“同時に”おこなうというので，Jacobi 反復のことを**同時緩和法**(simultaneous relaxation)と呼ぶこともある．

方程式系の係数行列 $A=[a_{ij}]$ を，その対角部分 $D=\mathrm{diag}[a_{11},\cdots,a_{nn}]$，下三角部分 $L=[l_{ij}]$ ($l_{ij}=a_{ij}$, $j<i$；$l_{ij}=0$, $j\geqq i$)，上三角部分 $U=[u_{ij}]$ ($u_{ij}=0$, $j\leqq i$；$u_{ij}=a_{ij}$, $j>i$) の和に表す：

$$A=L+D+U. \tag{4.4.15}$$

(ここでの D,L,U は，2.15 節で論じた A の LU 分解，あるいは LDU 分解の D,L,U とは異なることに注意！)

すると，Jacobi 反復は

$$\boldsymbol{x}^{(\nu+1)}=D^{-1}(\boldsymbol{b}-(L+U)\boldsymbol{x}^{(\nu)}) \tag{4.4.16}$$

と書けるが，(4.4.3)～(4.4.5) の形式に表すときにはこれが

$$\boldsymbol{d}^{(\nu)}=B\boldsymbol{r}^{(\nu)}, \qquad B=D^{-1} \tag{4.4.17}$$

とおいたものに相当することは明らかであろう．(4.4.7) の近似誤差 $\boldsymbol{\varepsilon}^{(\nu)}$ は

$$\boldsymbol{\varepsilon}^{(\nu+1)}=-D^{-1}(L+U)\boldsymbol{\varepsilon}^{(\nu)} \tag{4.4.18}$$

を満たす．

Jacobi 反復が収束するための十分条件として，係数行列 A が**対角優位**(diagonally dominant)であることという条件が知られている．これは以下のように容易に示せる．

n 次正方行列 $A=[a_{ij}]$ は

$$|a_{ii}|>\sum_{j\neq i}|a_{ij}| \qquad (i=1,2,\cdots,n) \tag{4.4.19}$$

であるとき“行に関して”対角優位，

$$|a_{jj}|>\sum_{i\neq j}|a_{ij}| \qquad (j=1,2,\cdots,n) \tag{4.4.20}$$

であるとき“列に関して”対角優位であるという．

係数行列 A が行に関して対角優位であるときには

$$\mu=\max_i\sum_{j\neq i}|a_{ij}|/|a_{ii}| \tag{4.4.21}$$

とおく．もちろん $0\leqq\mu<1$ であるが，このとき

$$\|I-BA\|_\infty=\max_i\left|\frac{1}{|a_{ii}|}\sum_{j\neq i}|a_{ij}|\right|=\mu \tag{4.4.22}$$

であるから，$\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty \leqq \mu^\nu \|\boldsymbol{\varepsilon}^{(0)}\|_\infty$ で，$\boldsymbol{\varepsilon}^{(\nu)} \to \boldsymbol{0}\ (\nu \to \infty)$ となる（A が列に関して対角優位のときは $\|\cdot\|_1$ についてほぼ同様の議論ができる）．

注意 4.4.4 上の証明の副産物として「対角優位な行列は正則である」ということも示されたことになる．実際，$A\boldsymbol{x}=\boldsymbol{b}$ の解 $\boldsymbol{x}$ がこのような反復法で定められるということは，その解が一意に定まるということ，すなわち A が正則であることを意味している．

4.4.3 Gauss-Seidel 反復

Jacobi 反復において(4.4.13)により $x_i^{(\nu+1)}$ を定めるとき，右辺の $x_j^{(\nu)}$ のうち $x_1^{(\nu)}, \cdots, x_{i-1}^{(\nu)}$ をすでに値の更新されている $x_1^{(\nu+1)}, \cdots, x_{i-1}^{(\nu+1)}$ でおきかえることにすれば，直観的には“近似の改良速度が向上”しそうに思えよう．そのような方法が **Gauss-Seidel**（ガウス-ザイデル）**反復**（Gauss-Seidel iteration）である．すなわち，(4.4.12)の仮定のもとに，$i := 1, 2, \cdots, n$ に対して，

$$x_i^{(\nu+1)} = \frac{1}{a_{ii}}\left(b_i - \sum_{j<i} a_{ij} x_j^{(\nu+1)} - \sum_{j>i} a_{ij} x_j^{(\nu)}\right). \qquad (4.4.23)$$

係数行列を(4.4.15)の和の形に分解しておく．行列・ベクトル記法では

$$\boldsymbol{x}^{(\nu+1)} = D^{-1}(\boldsymbol{b} - L\boldsymbol{x}^{(\nu+1)} - U\boldsymbol{x}^{(\nu)}),$$

あるいは

$$\boldsymbol{x}^{(\nu+1)} = (D+L)^{-1}(\boldsymbol{b} - U\boldsymbol{x}^{(\nu)}) \qquad (4.4.24)$$

と表せる．(4.4.7)の近似誤差 $\boldsymbol{\varepsilon}^{(\nu)}$ は

$$\varepsilon_i^{(\nu+1)} = -\frac{1}{a_{ii}}\left(\sum_{j<i} a_{ij} \varepsilon_j^{(\nu+1)} + \sum_{j>i} a_{ij} \varepsilon_j^{(\nu)}\right) \qquad (i = 1, 2, \cdots, n), \qquad (4.4.25)$$

あるいは，同じことであるが，

$$\boldsymbol{\varepsilon}^{(\nu+1)} = -(D+L)^{-1} U \boldsymbol{\varepsilon}^{(\nu)} \qquad (4.4.26)$$

を満たす．

残差 $\boldsymbol{r}^{(\nu)}$ と修正量 $\boldsymbol{d}^{(\nu)}$ との関係は

$$\boldsymbol{d}^{(\nu)} = B\boldsymbol{r}^{(\nu)}, \qquad B = (D+L)^{-1} \qquad (4.4.27)$$

と表せる．

Gauss-Seidel 反復が収束するための十分条件として有用なのは次の二つであろう．

(i) 係数行列が対角優位であること．

(ii)　係数行列が正定値(Hermite)行列であること．

条件(ii)に関しては6.8.4項で扱うことにする．条件(i)が成立すれば収束列が得られることはJacobi反復のときと同様にして示せる．すなわち，(4.4.25)により，

$$
\begin{aligned}
|\varepsilon_1^{(\nu+1)}| &\leqq \frac{1}{|a_{11}|}\sum_{j=2}^{n}|a_{1j}||\varepsilon_j^{(\nu)}| \leqq \left(\frac{1}{|a_{11}|}\sum_{j=2}^{n}|a_{1j}|\right)\cdot\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty \leqq \mu\cdot\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty,\\
|\varepsilon_2^{(\nu+1)}| &\leqq \frac{1}{|a_{22}|}\left(|a_{21}|\cdot|\varepsilon_1^{(\nu+1)}| + \sum_{j=3}^{n}|a_{2j}|\cdot|\varepsilon_j^{(\nu)}|\right)\\
&\leqq \left(\frac{1}{|a_{22}|}\sum_{j\neq 2}|a_{2j}|\right)\cdot\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty \leqq \mu\cdot\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty,\\
&\cdots\\
&\cdots\\
|\varepsilon_n^{(\nu+1)}| &\leqq \cdots \leqq \mu\cdot\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty
\end{aligned}
$$

であるから

$$\|\boldsymbol{\varepsilon}^{(\nu+1)}\|_\infty \leqq \mu\cdot\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty. \tag{4.4.28}$$

ここで，μは(4.4.21)で定義した量で$0\leqq\mu<1$である．ゆえに，$\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty \leqq \mu^\nu\cdot\|\boldsymbol{\varepsilon}^{(0)}\|_\infty$で，$\nu\to\infty$のとき$\|\boldsymbol{\varepsilon}^{(\nu)}\|_\infty\to 0$．

注意 4.4.5　Jacobi反復では$x_1^{(\nu)},\cdots,x_n^{(\nu)}$から$x_1^{(\nu+1)},\cdots,x_n^{(\nu+1)}$を同時に計算して$\boldsymbol{x}^{(\nu)}$を$\boldsymbol{x}^{(\nu+1)}$に更新したが，Gauss-Seidel反復では，$x_1^{(\nu+1)},\cdots,x_n^{(\nu+1)}$を次々と計算し，計算し終わったものを直ちに$x_i^{(\nu)}$とおきかえて次の$x_i^{(\nu+1)}$の計算に用いるというやり方をするので，Gauss-Seidel反復のことを**逐次緩和法**(successive relaxation)と呼ぶこともある．また，計算に必要な作業領域の観点からすると，Jacobi反復では$\boldsymbol{x}^{(\nu+1)}$の全成分がすべて計算されるまで$\boldsymbol{x}^{(\nu)}$を記憶しておく必要があるが，Gauss-Seidel反復では$\boldsymbol{x}$の成分を1組だけ記憶しておいて，その中味を次々と更新して行けばよいという違いもある(Jacobi反復に比べてGauss-Seidel反復の方が作業領域が約半分で済む)．

4.4.4　その他の反復

緩和法においては，一般に，解への収束$\boldsymbol{x}^{(\nu)}\to\hat{\boldsymbol{x}}$は“1次収束”で，“加速”が有効である場合も少なくない．すなわち，**加速係数**(acceleration factor)ω $(1<\omega(\leqq 2))$を適当に選んで，(4.4.5)の代りに

$$\boldsymbol{x}^{(\nu+1)} = \boldsymbol{x}^{(\nu)} + \omega\boldsymbol{d}^{(\nu)} \tag{4.4.29}$$

とすると，収束が速くなることがある．このようなやり方は**過緩和法**(over-relaxation)と呼ばれている．

Jacobi 反復や Gauss-Seidel 反復において，第 $\nu+1$ 近似値を変数ごとに計算する代りに変数のブロックごとに計算することも可能である．たとえば，変数のベクトル $\boldsymbol{x}$ を s 個のブロック(同じ大きさである必要はない)に分けて

$$\boldsymbol{x}=[\boldsymbol{x}_1{}^{\top},\cdots,\boldsymbol{x}_s{}^{\top}]^{\top} \tag{4.4.30}$$

とし，係数行列もそれに対応して

$$A=\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1s} \\ A_{21} & A_{22} & \cdots & A_{2s} \\ \vdots & \vdots & \ddots & \vdots \\ A_{s1} & A_{s2} & \cdots & A_{ss} \end{bmatrix} \tag{4.4.31}$$

とブロック行列で表す(→2.4.1 項)．A_{ii} は正方行列であり，ここではみな正則であるとする．右辺ベクトル $\boldsymbol{b}$ も

$$\boldsymbol{b}=[\boldsymbol{b}_1{}^{\top},\cdots,\boldsymbol{b}_s{}^{\top}]^{\top} \tag{4.4.32}$$

と分割する．このとき**ブロック Jacobi 反復**(block Jacobi iteration)，**ブロック Gauss-Seidel 反復**(block Gauss-Seidel iteration)はそれぞれ次のような形となる．

$$\boldsymbol{x}_i^{(\nu+1)}=A_{ii}{}^{-1}\Bigl(\boldsymbol{b}_i-\sum_{j\neq i}A_{ij}\boldsymbol{x}_j{}^{(\nu)}\Bigr), \tag{4.4.33}$$

$$\boldsymbol{x}_i^{(\nu+1)}=A_{ii}{}^{-1}\Bigl(\boldsymbol{b}_i-\sum_{j<i}A_{ij}\boldsymbol{x}_j{}^{(\nu+1)}-\sum_{j>i}A_{ij}\boldsymbol{x}_j{}^{(\nu)}\Bigr). \tag{4.4.34}$$

有理解法(直接法)と反復解法との特徴(特長)を兼ね備えたものとして**共役勾配法**(conjugate gradient method；略して **CG 法**)およびその多数の変種があるが，本書の範囲を超えるので，ここでは省略する．

4.5　特殊な形の線形方程式系

特殊な形をした方程式系にはそれぞれ特別な解法がある．ここでは，ごく限られた 3 例についてだけ述べる．なお，演習問題 4.3 も見よ．

4.5.1　巡回行列を係数行列とする方程式系

方程式系 $A\boldsymbol{x}=\boldsymbol{b}$ の係数行列 A が n 次の巡回行列(→2.16.7項)である場合には，手間(必要な演算回数)が $n\log n$ に比例する程度で済む解法がある．実際，2.16.7項で述べたように，例2.14.7の Z で定義される**離散 Fourier 変換**を用いて

$$A=\begin{bmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-2} & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & \cdots & a_{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_2 & a_3 & \cdots & \cdots & a_0 & a_1 \\ a_1 & a_2 & \cdots & \cdots & a_{n-1} & a_0 \end{bmatrix}$$
$$=\sqrt{n}Z\cdot\mathrm{diag}[\alpha_0,\alpha_1,\cdots,\alpha_{n-1}]\cdot Z^{\mathrm{H}} \tag{4.5.1}$$

と表しておけば，

$$A\boldsymbol{x}=\sqrt{n}Z\cdot\mathrm{diag}[\alpha_0,\alpha_1,\cdots,\alpha_{n-1}]\cdot Z^{\mathrm{H}}\cdot\boldsymbol{x}=\boldsymbol{b} \tag{4.5.2}$$

より

$$\boldsymbol{x}=\frac{1}{\sqrt{n}}Z\cdot\mathrm{diag}[\alpha_0{}^{-1},\alpha_1{}^{-1},\cdots,\alpha_{n-1}{}^{-1}]\cdot Z^{\mathrm{H}}\cdot\boldsymbol{b} \tag{4.5.3}$$

という形で解が求められる．ここで，

$$\left.\begin{aligned} Z&=[z_{ij}], \quad z_{ij}=\frac{1}{\sqrt{n}}\exp(2\pi\mathrm{i}(i-1)(j-1)/n), \\ Z^{\mathrm{H}}&=[\bar{z}_{ji}], \quad \bar{z}_{ji}=\frac{1}{\sqrt{n}}\exp(-2\pi\mathrm{i}(i-1)(j-1)/n), \end{aligned}\right\} \tag{4.5.4}$$

$$\begin{bmatrix} \alpha_0 \\ \alpha_1 \\ \vdots \\ \alpha_{n-1} \end{bmatrix}=Z\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_{n-1} \end{bmatrix} \tag{4.5.5}$$

である．**高速 Fourier 変換**の算法を用いれば，Z や Z^{H} をベクトルに掛ける演算は $n\log n$ に比例する手間で遂行できるから，(4.5.3)，(4.5.5)により，巡回行列を係数とする線形方程式系は $n\log n$ に比例する手間で解くことができる．

4.5.2　Toeplitz 行列を係数行列とする方程式系

信号解析，時系列解析などにおいては

$$\begin{bmatrix} v^{(0)} & v^{(1)} & \cdots & v^{(n-2)} & v^{(n-1)} \\ v^{(1)} & v^{(0)} & \ddots & & v^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ v^{(n-2)} & & \ddots & \ddots & v^{(1)} \\ v^{(n-1)} & v^{(n-2)} & \cdots & v^{(1)} & v^{(0)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} = \begin{bmatrix} v^{(1)} \\ v^{(2)} \\ \vdots \\ v^{(n-1)} \\ v^{(n)} \end{bmatrix} \tag{4.5.6}$$

の形の方程式系を解く必要が頻繁に生じる(→2.16.6 項)．この形の方程式系に対しても最近では各種の“高速”解法が考案されており，$n \log n$ に比例する程度の手間の解法もあるが，ここでは古典的な n^2 に比例する程度の手間の解法(**Levinson-Durbin**(レヴィンソン-ダービン)**の解法**と呼ばれる方法)について述べる．これには，2.16.6 項の公式(2.16.53)～(2.16.55)を利用する．

$n=1,2,\cdots$と次数を増やしながら(4.5.6)の解を構成して行くので，次数が n のときの解を $x_k^{(n)}$ $(k=1,2,\cdots,n)$ と記すことにする．2.16.6 項の記法に従えば，(4.5.6)の係数行列を T_n として，Cramer の公式(→4.2 節)により，

$$x_k^{(n)} = \det T_n{}^k / \det T_n \qquad (k=1,\cdots,n) \tag{4.5.7}$$

となるので，(2.16.55′)により

$$x_k^{(n)} = x_k^{(n-1)} - x_{n-k}^{(n-1)} \cdot x_n^{(n)} \qquad (k=1,\cdots,n-1) \tag{4.5.8}$$

という n に関する漸化式が得られる．ここで，

$$x_n^{(n)} = \det T_n{}^n / \det T_n \tag{4.5.9}$$

は以下のようにして求められる．(4.5.6)の両辺に左から $\left[\begin{array}{c|c} (T_{n-1})^{-1} & O \\ \hline O & 1 \end{array}\right]$ を掛けると

$$\left[\begin{array}{ccc|c} 1 & & O & x_{n-1}^{(n-1)} \\ & \ddots & & \vdots \\ O & & 1 & x_1^{(n-1)} \\ \hline v^{(n-1)} & \cdots & v^{(1)} & v^{(0)} \end{array}\right] \left[\begin{array}{c} x_1^{(n)} \\ \vdots \\ x_{n-1}^{(n)} \\ \hline x_n^{(n)} \end{array}\right] = \left[\begin{array}{c} x_1^{(n-1)} \\ \vdots \\ x_{n-1}^{(n-1)} \\ \hline v^{(n)} \end{array}\right] \tag{4.5.10}$$

となる(T_{n-1} が行の順序と列の順序を同時に逆転しても不変であることに注意すれば，(4.5.10)の係数行列の最右列が $x_k^{(n-1)}$ $(k=n-1,n-2,\cdots,1)$ で表されることがわかる)．(4.5.10)にさらに左から $\left[\begin{array}{c|c} I_{n-1} & O \\ \hline -v^{(n-1)} \cdots -v^{(1)} & 1 \end{array}\right]$ を掛けると

$$\left[\begin{array}{ccc|c} 1 & & O & x_{n-1}^{(n-1)} \\ & \ddots & & \vdots \\ O & & 1 & x_1^{(n-1)} \\ \hline 0 & \cdots & 0 & w^{(n)} \end{array}\right]\begin{bmatrix} x_1^{(n)} \\ \vdots \\ x_{n-1}^{(n)} \\ \hline x_n^{(n)} \end{bmatrix}=\begin{bmatrix} x_1^{(n-1)} \\ \vdots \\ x_{n-1}^{(n-1)} \\ \hline u^{(n)} \end{bmatrix}, \tag{4.5.11}$$

$$u^{(n)}=v^{(n)}-\sum_{j=1}^{n-1}x_j^{(n-1)}v^{(n-j)}, \tag{4.5.12}$$

$$w^{(n)}=v^{(0)}-\sum_{j=1}^{n-1}x_{n-j}^{(n-1)}v^{(n-j)}, \tag{4.5.13}$$

が得られるから，$x_n^{(n)}$ は結局

$$x_n^{(n)}=u^{(n)}/w^{(n)} \tag{4.5.14}$$

で計算できる((4.5.8)は(4.5.11)からも導かれる)．

このように，$x_k^{(n-1)}(k=1,\cdots,n-1)$が知られているときに $x_k^{(n)}(k=1,\cdots,n)$を求めるには，まず(4.5.12)，(4.5.13)，(4.5.14)により $x_n^{(n)}$ を定めてから，(4.5.8)により $x_k^{(n)}(k=1,\cdots,n-1)$を定めればよい．この計算が，約 $3n$ 回の乗算と加減算と(それに1回の除算と)で遂行できることは明らかである．$n=1$ に対しては $x_1^{(1)}=v^{(1)}/v^{(0)}$ であるから，このやり方で $x_k^{(n)}(k=1,\cdots,n)$を求めるまでに必要な乗算と加減算の回数はそれぞれ約 $3n^2/2$ 回(それに除算 n 回)であり，一般の形の方程式を(たとえばCholesky分解法で)解く場合の $n^3/6$ 回に比べて n のオーダーが一つ少ない．

注意 4.5.1 (4.5.13)の $w^{(n)}$ の積和計算も倹約することができる．実際，(4.5.6)の係数行列 T_n を(4.5.11)の係数行列に変換するのに掛けた行列の行列式(の積)は$(\det T_{n-1})^{-1}$ であるから，

$$w^{(n)}=\det T_n/\det T_{n-1} \tag{4.5.15}$$

である．そこで，(2.16.53′)によれば，

$$\begin{aligned} w^{(n)}&=[1-(\det T_{n-1}{}^{n-1}/\det T_{n-1})^2]\,w^{(n-1)} \\ &=(1-(x_{n-1}^{(n-1)})^2)\,w^{(n-1)}, \end{aligned} \tag{4.5.16}$$

したがって，(4.5.13)を用いずに，直接(4.5.16)により $w^{(1)}(=v^{(0)}),w^{(2)},\cdots$を順に計算できる．

注意 4.5.2 上のようにして(4.5.6)の解 $x_k^{(n)}$ を構成して行くのと並行して，右辺を一般のベクトルとした方程式

$$\begin{bmatrix} v^{(0)} & v^{(1)} & \cdots & \cdots & v^{(n-1)} \\ v^{(1)} & v^{(0)} & \ddots & & v^{(n-2)} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & v^{(1)} \\ v^{(n-1)} & v^{(n-2)} & \cdots & v^{(1)} & v^{(0)} \end{bmatrix} \begin{bmatrix} y_1^{(n)} \\ y_2^{(n)} \\ \vdots \\ \vdots \\ y_n^{(n)} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ \vdots \\ b_n \end{bmatrix} \tag{4.5.17}$$

の解を構成して行くことができる．実際，上と同様にして，(4.5.17)に左から $\left[\begin{array}{c|c} (T_{n-1})^{-1} & O \\ \hline [-v^{(n-1)},\cdots,-v^{(1)}](T_{n-1})^{-1} & 1 \end{array}\right]$ を掛けると，

$$\left[\begin{array}{ccc|c} 1 & & O & x_{n-1}^{(n-1)} \\ & \ddots & & \vdots \\ O & & 1 & x_1^{(n-1)} \\ \hline & O & & w^{(n)} \end{array}\right] \begin{bmatrix} y_1^{(n)} \\ \vdots \\ y_{n-1}^{(n)} \\ \hline y_n^{(n)} \end{bmatrix} = \begin{bmatrix} y_1^{(n-1)} \\ \vdots \\ y_{n-1}^{(n-1)} \\ \hline s^{(n)} \end{bmatrix}, \tag{4.5.18}$$

$$s^{(n)} = b_n - \sum_{j=1}^{n-1} y_j^{(n-1)} v^{(n-j)} \tag{4.5.19}$$

($w^{(n)}$ は(4.5.13)，(4.5.16)と同じ)を得るので，$y_k^{(n-1)}$ $(k=1,\cdots,n-1)$ が知られていれば，

$$y_n^{(n)} = s^{(n)}/w^{(n)}, \tag{4.5.20}$$

$$y_k^{(n)} = y_k^{(n-1)} - x_{n-k}^{(n-1)} \cdot y_n^{(n)} \qquad (k=1,\cdots,n-1) \tag{4.5.21}$$

によって $y_k^{(n)}$ $(k=1,\cdots,n)$ を定めることができる．

4.5.3 Sylvester の方程式

m 次正方行列 $A=[a_{ik}]$，n 次正方行列 $B=[b_{lj}]$，$m\times n$ 型行列 $C=[c_{ij}]$ が与えられたとき，

$$AX - XB = C \tag{4.5.22}$$

を満たす $m\times n$ 型行列 $X=[x_{kl}]$(の要素)を求める問題を考える$(i,k=1,\cdots,m\,;\,l,j=1,\cdots,n)$．(4.5.22)は **Sylvester の方程式**と呼ばれる．

Sylvester の方程式は X の $m\cdot n$ 個の要素 x_{kl} を一列に並べたものを未知数ベクトル(それを $\mathscr{X}$ と記す)とした $m\cdot n$ 次元の線形方程式系とみなすことができる．方程式も $m\cdot n$ 個あり，右辺ベクトルは C の要素を一列に並べたもの($\mathscr{C}$ と記す)となる．このとき，右辺が c_{ij} である方程式の中で未知数 x_{kl} にかかる係数は

$$a_{ik}\delta_{lj} - \delta_{ik} b_{lj} \tag{4.5.23}$$

である($i\neq j$ のとき $\delta_{ij}=0$；$i=j$ のとき $\delta_{ij}=1$)．以下では，未知数も方程式も

(k,l), (i,j)を$(1,n)$, $(1,n-1)$, …, $(1,2)$, $(1,1)$；$(2,n)$, $(2,n-1)$, …, $(2,2)$, $(2,1)$；…；(m,n), $(m,n-1)$, …, $(m,2)$, $(m,1)$という順序(いわゆる“辞書式順序”の一種)に並べて議論を進めることにする．このようにして作られる線形方程式系を

$$\mathscr{A}\mathscr{X}=\mathscr{C} \tag{4.5.24}$$

と表しておく($\mathscr{A}$ の要素は(4.5.23))．

まず特別な場合として，$A=[a_{ik}]$も $B=[b_{lj}]$も上三角行列である場合を扱う．このときには，$i>k$ なら $a_{ik}=0$, $l>j$ なら $b_{lj}=0$ であるから，上記の順序で(i,j)が(k,l)より後にある(すなわち，$i>k$ であるか，あるいは $i=k$ で $l>j$ であるか)であれば(4.5.23)において $a_{ik}=\delta_{ik}=0$ あるいは $\delta_{lj}=b_{lj}=0$，したがって $a_{ik}\delta_{lj}-\delta_{ik}b_{lj}=0$ となる．このことは，$\mathscr{A}$ が“上三角行列”であることを意味している．そこで，$\mathscr{A}$ の“対角要素”($(i,j)=(k,l)$)がみな非零，すなわち

$$a_{ii}-b_{jj}\neq 0 \qquad (i=1,\cdots,m\,;\ j=1,\cdots,n) \tag{4.5.25}$$

であれば，方程式系(4.5.24)は，x_{kl} について(k,l)を一列に並べた順序に従って順々に解いて行くことができる(少なくとも一つの(i,j)に対して $a_{ii}=b_{jj}$ であれば，$\det\mathscr{A}=0$ で，方程式(4.5.22)には解が存在しないか，多数存在するかの，どちらかである)．

後に5.2.1項で示すように，適当な n 次正則行列 S と m 次正則行列 T によって SAS^{-1} と TBT^{-1} を共に上三角にすることができる(S,T をユニタリ行列に限ってもこうすることが可能である)から，そのような S,T を用いると(4.5.22)は

$$SAS^{-1}SXT^{-1}-SXT^{-1}TBT^{-1}=SCT^{-1} \tag{4.5.26}$$

と書き直せる．$Y=SXT^{-1}$ とおけば，(4.5.26)は，Y を未知数とする，上で調べた特別な形の Sylvester の方程式であるから，Y について直ちに解くことができ，$X=S^{-1}YT$ によって X を定めることができる．

注意 4.5.3　5.2.1項で示す通り，A,B を上のように上三角化したときにそれぞれの対角要素には A の固有値(重複度も数えれば m 個ある)と B の固有値(同じく n 個ある)が並ぶので，Sylvester の方程式の一意可解性の条件(4.5.25)は「A の固有値はどれも B の固有値と一致しないこと」と，より一般的に述べることができる．

注意 4.5.4　Sylvester の方程式は数学，特に線形代数の中でもよく利用される

が，応用上も制御系の解析や設計に役に立つものである．

4.6 最小二乗法

簡単のため $K=\mathbb{R}$ とする．ある変量 y がいくつかの変量 $x_1,\cdots,x_m$ の 1 次関数

$$y=a_0+a_1x_1+\cdots+a_mx_m \tag{4.6.1}$$

であるらしいと思って(これを“モデル”と呼ぶ)，多くの“観測”あるいは“測定”をしてデータ

$$(x_{1k},\cdots,x_{mk},y_k) \qquad (k=1,\cdots,N) \tag{4.6.2}$$

を得たとする．このとき，これらのデータが(4.6.1)を“なるべく良く満たす”ように 1 次式の係数 $a_0,a_1,\cdots,a_m$ を決定する問題を考える．“なるべく良く満たす”ということの意味を，“データを x_i,y に代入したときにモデルの右辺と左辺の喰い違い

$$\varepsilon_k=(a_0+a_1x_{1k}+\cdots+a_mx_{mk})-y_k \tag{4.6.3}$$

が，$\boldsymbol{\varepsilon}=[\varepsilon_1,\cdots,\varepsilon_N]^{\mathsf{T}}$ の 2 乗ノルム(の 2 乗)

$$\|\boldsymbol{\varepsilon}\|_2{}^2=\sum_{k=1}^{N}|\varepsilon_k|^2 \tag{4.6.4}$$

が最小になるという意味で，“最も小さい”ということであると定義する．このようにして，$a_0,\cdots,a_m$ を定める方法を**最小二乗法**(method of least squares)という．(ここでは，x_{ik} や y_k が所与量(データ)で，$a_0,a_1,\cdots,a_m$ が決定さるべき未知量であることに注意！)

データ行列(data matrix)を

$$[A|\boldsymbol{b}]=\left[\begin{array}{ccccc|c} 1 & x_{11} & x_{21} & \cdots & x_{m1} & y_1 \\ 1 & x_{12} & x_{22} & \cdots & x_{m2} & y_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 1 & x_{1N} & x_{2N} & \cdots & x_{mN} & y_N \end{array}\right] \tag{4.6.5}$$

で定義し，$\boldsymbol{b}=[y_1,\cdots,y_N]^{\mathsf{T}}$，$\boldsymbol{\alpha}=[a_0,a_1,\cdots,a_m]^{\mathsf{T}}$ とおけば，

$$\boldsymbol{\varepsilon}=A\boldsymbol{\alpha}-\boldsymbol{b} \tag{4.6.6}$$

と書かれるから，

$$\|\boldsymbol{\varepsilon}\|_2^2 = \|A\boldsymbol{\alpha} - \boldsymbol{b}\|_2^2 = \boldsymbol{\alpha}^\top A^\top A\boldsymbol{\alpha} - 2\boldsymbol{\alpha}^\top A^\top \boldsymbol{b} + \boldsymbol{b}^\top \boldsymbol{b} \tag{4.6.7}$$

である．(4.6.7)は $\boldsymbol{\alpha}$ の(非負定値)二次形式と線形形式と定数項とよりなるから，その最小値を与える $\boldsymbol{\alpha}$ は，右辺を a_i $(i=0,\cdots,m)$で微分して 0 とおくことにより，

$$(A^\top A)\boldsymbol{\alpha} = A^\top \boldsymbol{b} \tag{4.6.8}$$

を満たさなければならない．(4.6.8)は $\boldsymbol{\alpha}$ を未知ベクトルとする線形方程式系であり，$\boldsymbol{\alpha}$ を決定するための**正規方程式**(normal equation)と呼ばれる．この係数行列 $A^\top A$ は明らかに非負定値であり，“モデル式”(4.6.1)に“無駄がなく”かつ“独立な観測量が十分多くあれば”正定値である．

正規方程式を経由して $\boldsymbol{\alpha}$ を定めるのが最小二乗法の古典的解法であるが，データ行列の QR 分解(→2.15.5 項)を用いると，より直接的に $\boldsymbol{\alpha}$ を定めることもできる[前者でも後者でも必要とされる計算の手間はほぼ等しいとか，後者の方が丸め誤差に強いとかいうような話もあるが，詳しくは省く]．すなわち，2 乗ノルムが直交変換により不変であること($\|\boldsymbol{\varepsilon}\|_2 = \|Q^\top \boldsymbol{\varepsilon}\|_2$，$Q$：直交行列)を利用して，行列 A を

$$A = QR \qquad (Q：N \text{ 次直交行列}，R：N\times(m+1) \text{ 型上三角行列}) \tag{4.6.9}$$

と QR 分解し，

$$\tilde{\boldsymbol{b}} = Q^\top \boldsymbol{b} \tag{4.6.10}$$

とおけば，(4.6.6)の両辺に $Q^\top$ を左から掛けることにより

$$\tilde{\boldsymbol{\varepsilon}} = Q^\top \boldsymbol{\varepsilon} = R\boldsymbol{\alpha} - \tilde{\boldsymbol{b}} \tag{4.6.11}$$

を得る．(4.6.11)は

$$\begin{bmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \vdots \\ \vdots \\ \vdots \\ \tilde{\varepsilon}_N \end{bmatrix} = \begin{bmatrix} r_{00} & & \boldsymbol{r_{ij}} \\ & r_{11} & \\ & & \ddots \\ & & & r_{mm} \\ & O & \\ & & \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} - \begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \vdots \\ \vdots \\ \tilde{b}_N \end{bmatrix} \tag{4.6.12}$$

という形をしている．この形から直ちに，rank $A = m+1$ のとき，

(i)　$\|\tilde{\boldsymbol{\varepsilon}}\|_2^2$ を小さくしようとすれば，$\tilde{\varepsilon}_1 = \cdots = \tilde{\varepsilon}_{m+1} = 0$ になるように $\boldsymbol{\alpha}$ を定めることができる，

(ii)　$\boldsymbol{\alpha}$ をどのように定めても，$\tilde{\varepsilon}_{m+j}=\tilde{b}_{m+j}$　$(j=2,\cdots,N-m)$ である，ということがわかる．すなわち，$\|\tilde{\boldsymbol{\varepsilon}}\|_2$ の最小値は，$a_0,a_1,\cdots,a_m$ を

$$\begin{bmatrix} r_{00} & & \boldsymbol{r_{ij}} \\ & r_{11} & \\ & & \ddots & \\ O & & & r_{mm} \end{bmatrix}\begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_m \end{bmatrix}=\begin{bmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ \vdots \\ \tilde{b}_{m+1} \end{bmatrix} \tag{4.6.13}$$

により定めれば達成され，そのとき

$$\min\|\tilde{\boldsymbol{\varepsilon}}\|_2{}^2(=\min\|\boldsymbol{\varepsilon}\|_2{}^2)=|\tilde{b}_{m+2}|^2+\cdots+|\tilde{b}_N|^2 \tag{4.6.14}$$

である．

注意 4.6.1　モデル(4.6.1)が1次式であることから，この種の最小二乗法のことを“線形”最小二乗法を呼ぶこともある．もちろん，自然に，“非線形”最小二乗法への拡張は可能であり，その方が多くの応用があるが，それについては本書では扱わない．

演 習 問 題

4.1　次の線形方程式系をいろいろな方法で解け．

$$\begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}=\begin{bmatrix} 3 \\ 2 \\ 2 \\ 3 \end{bmatrix}$$

4.2　次の線形方程式系を解け．

$$\begin{bmatrix} 2 & -1 & & & & \\ -1 & 2 & -1 & & O & \\ & -1 & 2 & -1 & & \\ & & \ddots & \ddots & \ddots & \\ & O & & -1 & 2 & -1 \\ & & & & -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix}=\begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \\ 1 \end{bmatrix}$$

[略解]　このような形の方程式の場合，x_i は線形定数係数差分方程式 $-x_{i+1}+2x_i-x_{i-1}=1\,(i=1,\cdots,n)$ の境界条件 $x_0=0, x_{n+1}=0$ の下での解として定めることができる．$x_i=-i^2/2$ が“特解”であるから，解を $x_i=-i^2/2+y_i$ とおけば，y_i は同次差分方程式 $-y_{i+1}+2y_i-y_{i-1}=0$ と境界条件 $y_0=0, y_{n+1}=(n+1)^2/2$ を満たす．この差分方程式の特性方程式は $-\lambda^2+2\lambda-1=-(\lambda-1)^2=0$ で，その解は $\lambda_1=1$(2重根)であるから，同次方程式の一般解は $y_i=c_1\lambda_1{}^i+c_2 i\lambda_1{}^i=c_1+c_2 i$ と表せる(c_1, c_2 は定数)．境界条

件を満たすように c_1, c_2 を定めると，$c_1=0, c_2=(n+1)/2$．ゆえに，もとの方程式の解は $x_i=i(n+1-i)/2$．

4.3 A が正則で，$A\boldsymbol{x}=\boldsymbol{b}$ の解 $\boldsymbol{x}$ が得られているとき，$\tilde{A}\tilde{\boldsymbol{y}}=\tilde{\boldsymbol{b}}$ の解を求めよ．ただし，

$$A=\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}, \quad \boldsymbol{b}=\begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \quad \boldsymbol{x}=\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix};$$

$$\tilde{A}=\left[\begin{array}{c|c} & a_{1,n+1} \\ A & \vdots \\ & a_{n,n+1} \\ \hline a_{n+1,1}\cdots a_{n+1,n} & a_{n+1,n+1} \end{array}\right], \quad \tilde{\boldsymbol{b}}=\left[\begin{array}{c} \boldsymbol{b} \\ \hline b_{n+1} \end{array}\right], \quad \tilde{\boldsymbol{y}}=\left[\begin{array}{c} y_1 \\ \vdots \\ y_n \\ \hline y_{n+1} \end{array}\right]$$

［略解］ $\boldsymbol{x}=A^{-1}\boldsymbol{b}$ であるから，$\boldsymbol{a}=[a_{1,n+1},\cdots a_{n,n+1}]^{\top}, \boldsymbol{a}'=[a_{n+1,1}\cdots a_{n+1,n}]$ とおき，$\tilde{A}\tilde{\boldsymbol{y}}=\tilde{\boldsymbol{b}}$ に左から $\left[\begin{array}{c|c} A^{-1} & O \\ \hline -\boldsymbol{a}'A^{-1} & 1 \end{array}\right]$ を掛けると，$\left[\begin{array}{c|c} I_n & A^{-1}\boldsymbol{a} \\ \hline O & w \end{array}\right]\tilde{\boldsymbol{y}}=\left[\begin{array}{c} \boldsymbol{x} \\ \hline u \end{array}\right]$ ($w=a_{n+1,n+1}-\boldsymbol{a}'A^{-1}\boldsymbol{a}$, $u=b_{n+1}-\boldsymbol{a}'\cdot\boldsymbol{x}$) を得る．したがって，$w\neq 0$ ならば，$\tilde{\boldsymbol{y}}=[\boldsymbol{y}^{\top}, y_{n+1}]^{\top}$ とおけば，$y_{n+1}=u/w, \boldsymbol{y}=\boldsymbol{x}-y_{n+1}(A^{-1}\boldsymbol{a})$．［注：このように順次に次数を拡大しながら解を構成して行く方法を **escalator 法**などと呼んだこともあったが，本質的には，公式(2.10.8)と同内容のものであるともみなせるし，また A の LU 分解による方法の手順を順に追って行くのと同じことであるともみなせる．］

4.4 Sylvester の方程式(4.5.22)を線形方程式系の通常の形(4.5.24)に書いたときの係数行列 $\mathcal{A}$ を Kronecker 積(→2.4.2 項)を用いて表してみよ．

［略解］ 具体的に(4.5.23)を書き下してみる．$B'=[b_{ij}']$ を $b_{ij}'=b_{n+1-i,n+1-j}$ で定義し，(4.5.23)で導入した“辞書的順序”に従えば，$\mathcal{A}=A\otimes I_n-I_m\otimes B'$ であることがわかる．

5
固　有　値

本章では簡単のため $K=\mathbb{C}$ とする．議論のかなりの部分は，K が一般の体の場合でも成立するし，特に K が“代数的閉体”の場合にはほとんどの議論が成立する．

特に断らない限り，線形写像 $A: V_n \to V_n$ あるいはその表現行列である n 次正方行列 A について論ずる(一つの n 次元ベクトル空間のそれ自身への写像であることに注意)．また，“ユニタリ行列”というときには，V_n が計量空間であって，“正規直交基底”を採用して行列表現がなされているとする．すなわち，そのベクトルの長さ(2 乗ノルム)がユニタリ変換に関して不変であるようにしてあるものとする．V_n に計量が指定されていないときには，それにたまたま選んだ基底が“正規直交系”をなすような計量(“単位計量”)を仮に導入して，V_n を計量空間とみなしているものとする．しかし，固有値は計量とは本来無関係なので，計量が指定されていない場合のユニタリ変換(ユニタリ行列)は応用上はかなり便宜的な手段であると考えるべきである．

5.1　固有値と固有ベクトル

n 次正方行列 A に対してあるベクトル $\boldsymbol{x}$ ($\in V_n$, $\boldsymbol{x}\neq\boldsymbol{0}$) とスカラー λ ($\in\mathbb{C}$) が存在して

$$A\boldsymbol{x}=\lambda\boldsymbol{x} \tag{5.1.1}$$

を満たすとき，λ を A の一つの**固有値**(eigenvalue, proper value, characteristic value)，$\boldsymbol{x}$ を λ に対応する一つの**固有ベクトル**(eigenvector，など)と呼ぶ．

注意 5.1.1　$\boldsymbol{x}$ が A の固有値 λ に対応する固有ベクトルなら，$\boldsymbol{x}$ の任意のスカ

ラー倍 $\boldsymbol{y}=\alpha\boldsymbol{x}$ ($\alpha\in\mathbb{C}, \alpha\neq 0$) もそうであるから，固有ベクトルは“方向”だけが意味をもつものである．そこで，固有ベクトル $\boldsymbol{x}$ は単位長に“規格化”して扱うことも多い：$\|\boldsymbol{x}\|_2=1$.

注意 5.1.2 (5.1.1)で定義される $\boldsymbol{x}$ を**右固有ベクトル** (right eigenvector，など) と呼び，

$$\boldsymbol{y}^{\mathrm{H}}A=\lambda\boldsymbol{y}^{\mathrm{H}} \quad \text{あるいは} \quad A^{\mathrm{H}}\boldsymbol{y}=\bar{\lambda}\boldsymbol{y} \tag{5.1.2}$$

を満たす $\boldsymbol{y}$ ($\in V_n$, $\boldsymbol{y}\neq\boldsymbol{0}$) を**左固有ベクトル** (left eigenvector，など) と呼ぶこともある．後者は A の Hermite 共役 A^{H} (2.3.1 項) の右固有ベクトルである (対応する固有値は複素共役となる)．

ベクトル空間 V_n の基底変換 (正則行列 S で表現されるとする) に伴って，A は相似変換を受けて $A'=S^{-1}AS$ となるが，(5.1.1) に左から S^{-1} を掛け，$\boldsymbol{x}=S\boldsymbol{y}$ とおけば

$$S^{-1}AS\boldsymbol{y}=\lambda S^{-1}\boldsymbol{x} \quad \text{すなわち} \quad A'\boldsymbol{y}=\lambda\boldsymbol{y} \tag{5.1.3}$$

となり，A' は A と同じ固有値をもち，対応する固有ベクトルは基底変換に伴って変換されるだけであることがわかる．

(5.1.1) を

$$(A-\lambda I)\boldsymbol{x}=\boldsymbol{0} \tag{5.1.4}$$

と書いてみれば，$A-\lambda I$ を係数行列とし，右辺が $\boldsymbol{0}$ である線形方程式系を得るから，これが $\boldsymbol{0}$ でない解 $\boldsymbol{x}$ をもつためには係数行列は特異でなければならない (→ 4.3.2 項)．(5.1.4) の係数行列 $A-\lambda I$ の各要素を変数 λ の 1 次式とみなせば，$\det(A-\lambda I)$ は λ の n 次多項式である．この多項式

$$\varphi_A(\lambda)=\det(A-\lambda I) \tag{5.1.5}$$

を A の**特性多項式** (characteristic polynomial) と呼ぶ (A と A' が相似なら $\varphi_A(\lambda)=\varphi_{A'}(\lambda)$ である (→ 注意 5.1.6)．また $\varphi_{A^{\mathrm{T}}}(\lambda)=\varphi_A(\lambda)$ も行列式の定義から明らか)．

λ が A の固有値であるためには，λ が特性多項式の零点でなければならない：

$$\varphi_A(\lambda)=0. \tag{5.1.6}$$

代数方程式 (5.1.6) は，A の**特性方程式** (characteristic equation) と呼ばれる (A が Hermite 行列の場合に**永年方程式** (secular equation) と呼ばれたこともある)．

逆に，λ が特性多項式の零点であれば $\det(A-\lambda I)=0$，すなわち $\mathrm{rank}(A-\lambda I)<n$ であるから，線形方程式系 (5.1.4) には $\boldsymbol{0}$ でない解 $\boldsymbol{x}$ がある．このよ

うな $\lambda, \boldsymbol{x}$ は(5.1.1)を満たすから，λ は A の固有値であることがわかる．

$\varphi_A(\lambda)$ が n 次多項式であるから，多重度まで含めて数えれば，n 次の正方行列 A の固有値は n 個存在する．

A の固有値の集合 $\{\lambda_1, \cdots, \lambda_n\}$ $(\varphi_A(\lambda_i)=0)$ のことを A の**スペクトル**(spectrum)と呼び，sp A と記すこともある(6.8.5 項の定義とも比較せよ)．

A の固有値の絶対値の最大なものの絶対値を $\rho(A)$ と記し，A の**スペクトル半径**(spectral radius)という．

注意 5.1.3　A のスペクトル半径は A の自然ノルム(→ 3.3.2 項)を超えない：

$$\rho(A) \leqq \|A\|. \tag{5.1.7}$$

実際，$\rho(A)=|\lambda_1|=\max_i |\lambda_i|$ とし，λ_1 に対応する固有ベクトルを $\boldsymbol{x}_1$ とすると，

$$\rho(A)=\frac{\|A\boldsymbol{x}_1\|}{\|\boldsymbol{x}_1\|} \leqq \max_x \frac{\|A\boldsymbol{x}\|}{\|\boldsymbol{x}\|}=\|A\|. \tag{5.1.8}$$

注意 6.8.3 も見よ．

注意 5.1.4　以上の議論によって，固有値が(多重度まで数えれば) n 個あり，一つの固有値の値に対して少なくとも 1 本の固有ベクトルが存在することはいえる．しかし，固有値が n 個あるからといって，n 個の独立な固有ベクトルが存在するわけではない．ただ，“異なる”固有値 $\{\lambda_1, \cdots, \lambda_m\}$ に対応する固有ベクトル $\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_m\}$ は線形独立である．実際，$\sum_{j=1}^{m} \alpha_j \boldsymbol{x}_j=\boldsymbol{0}$ に $I, A, A^2, \cdots, A^{m-1}$ を掛けると

$$\sum_{j=1}^{m} \lambda_j{}^k \alpha_j \boldsymbol{x}_j=\boldsymbol{0} \qquad (k=0,1,\cdots,m-1) \tag{5.1.9}$$

を得るが，$m \times m$ 型行列 $C=[\lambda_j{}^i]$ は Vandermonde 行列で，λ_j $(j=1,\cdots,m)$ が互いに異なる値のときには正則である(→ 2.16.1 項)から，(5.1.9)より $\alpha_j \boldsymbol{x}_j=\boldsymbol{0}$ $(j=1,\cdots,m)$ を得る．$\boldsymbol{x}_j \neq \boldsymbol{0}$ としてあるから，$\alpha_j=0$ $(j=1,\cdots,m)$．

A が n 本の線形独立な固有ベクトルの組 $B=\{\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n\}$ (対応する固有値を $\lambda_1, \cdots, \lambda_n$ とする)をもてば，B を V_n の基底に選ぶと，A の表現行列は明らかに $\Lambda=\mathrm{diag}[\lambda_1, \cdots, \lambda_n]$ である．言いかえれば，$\boldsymbol{x}_j$ を列 j にもつ行列を X とすると，

$$AX=X\Lambda \quad \text{すなわち} \quad X^{-1}AX=\Lambda. \tag{5.1.10}$$

この場合，A は固有値を対角要素とする対角行列に相似である．

注意 5.1.5　特性多項式を視覚に訴えるように表示すると

$$\varphi_A(\lambda)=\det\begin{bmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{bmatrix} \tag{5.1.11}$$

となるが，行列式を定義式(2.6.1)に従って $(-\lambda)$ の多項式に展開してみればわか

るように，

$$\varphi_A(\lambda)=(-\lambda)^n+c_{n-1}(-\lambda)^{n-1}+\cdots+c_k(-\lambda)^k+\cdots+c_1(-\lambda)+c_0 \qquad (5.1.12)$$

の係数 $c_n=1, c_{n-1}, \cdots, c_1, c_0$ は

$$c_{n-k}=A \text{ の } k \text{ 次の主小行列式の和} \qquad (5.1.13)$$

である．2.7.2，3 項の記号を用いれば

$$c_{n-k}=\mathrm{tr}\, A^{<k>} \qquad (5.1.13')$$

と表せる．（A の k 次の主小行列式 $|A|_{i_1\cdots i_k}^{i_1\cdots i_k}$ （$A^{<k>}$ の対角要素)(→ 2.7.1 項)は全部で $\binom{n}{k}$ 個ある．）特に，

$$c_{n-1}=\mathrm{tr}\, A, \qquad c_0=\det A \qquad (5.1.14)$$

である．一方，A の n 個の固有値(重複度も含めて)を $\lambda_1, \cdots, \lambda_n$ とすれば，

$$\varphi_A(\lambda)=(\lambda_1-\lambda)(\lambda_2-\lambda)\cdots(\lambda_n-\lambda) \qquad (5.1.15)$$

で，Vieta の関係式

$$c_{n-k}=\lambda_1, \cdots, \lambda_n \text{ の } k \text{ 次の基本対称式} \qquad (5.1.16)$$

によって，A の“k 次の主小行列式の和”と A の“固有値の k 次の基本対称式”とが等しいことがわかる．特に，

$$\mathrm{tr}\, A=\sum_{i=1}^{n}\lambda_i, \qquad \det A=\prod_{i=1}^{n}\lambda_i. \qquad (5.1.17)$$

(5.1.17)の第 2 式は，「A が正則であるための必要十分条件は固有値 0 をもたないことである」ということも表している．

注意 5.1.6 相似変換により固有値は変化しないから，特性多項式も変化しない．このことは直接

$$\begin{aligned}\varphi_{S^{-1}AS}(x)&=\det(S^{-1}AS-xI)=\det(S^{-1}(A-xI)S)\\&=(\det S^{-1})(\det(A-xI))(\det S)=\det(A-xI)\\&=\varphi_A(x)\end{aligned} \qquad (5.1.18)$$

としても確かめられる．

5.2 Schur 形と応用

5.2.1 Schur 形

n 次正方行列 A はユニタリ相似変換(＝ユニタリ合同変換)によって固有値を対角要素にもつ上三角行列に変換することができる．すなわち，A の固有値を λ_i （$i=1, \cdots, n$)とすると，

$$Q^{\mathrm{H}}AQ=R \qquad (5.2.1)$$

($R=[r_{ij}]$ は上三角行列で，$r_{ii}=\lambda_i$)を満たすようなユニタリ行列 Q が存在する．

この R を A の **Schur 形**という(A に対して Schur 形が一意に定まるわけではない).

この事実は，以下のように，“構成的に”示すことができる.

A の固有値の一つ λ_1 とそれに対応する固有ベクトル $\boldsymbol{u}_1$ を選ぶ ($\|\boldsymbol{u}_1\|_2=1$ と規格化しておく). A を表現している基底$\{\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$に $\boldsymbol{u}_1$ を付け加えたもの$\{\boldsymbol{u}_1,\boldsymbol{e}_1,\cdots,\boldsymbol{e}_n\}$から，(たとえば 3.6 節の Gram-Schmidt の直交化法により) $\boldsymbol{u}_1$ を第 1 要素とする正規直交系$\{\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n\}$を作ると，$Q^{(1)}\equiv[\boldsymbol{u}_1,\boldsymbol{u}_2,\cdots,\boldsymbol{u}_n]$はユニタリ行列であり，$A\boldsymbol{u}_1=\lambda_1\boldsymbol{u}_1$ であるから，

$$AQ^{(1)}=Q^{(1)}\left[\begin{array}{c|c}\lambda_1 & \ast \\ \hline O & A^{(1)}\end{array}\right],$$

すなわち

$$Q^{(1)\mathrm{H}}AQ^{(1)}=\left[\begin{array}{c|c}\lambda_1 & \ast \\ \hline O & A^{(1)}\end{array}\right] \tag{5.2.2}$$

となることがわかる．同様のことを $n-1$ 次の行列 $A^{(1)}$ に対しておこなうと，$n-1$ 次のユニタリ行列 Q' で

$$Q'^{\mathrm{H}}A^{(1)}Q'=\left[\begin{array}{c|c}\lambda_2 & \ast \\ \hline O & A^{(2)}\end{array}\right] \tag{5.2.3}$$

を満たすものが得られる.

$$Q^{(2)}=\left[\begin{array}{c|c}1 & O \\ \hline O & Q'\end{array}\right] \tag{5.2.4}$$

とすれば

$$\begin{aligned}Q^{(2)\mathrm{H}}Q^{(1)\mathrm{H}}AQ^{(1)}Q^{(2)}&=(Q^{(1)}Q^{(2)})^{\mathrm{H}}A(Q^{(1)}Q^{(2)})\\&=\left[\begin{array}{cc|c}\lambda_1 & \ast & \ast \\ 0 & \lambda_2 & \ast \\ \hline \multicolumn{2}{c|}{O} & A^{(2)}\end{array}\right]\end{aligned} \tag{5.2.5}$$

を得る. λ_2 は $A^{(1)}$ の固有値であるが A の固有値でもあることは(5.2.5)の形からわかる：

$$\begin{aligned}\varphi_A(\lambda)&=(\lambda_1-\lambda)\,\varphi_{A^{(1)}}(\lambda)\\&=(\lambda_1-\lambda)\,(\lambda_2-\lambda)\,\varphi_{A^{(2)}}(\lambda)\,.\end{aligned} \tag{5.2.6}$$

同様なことをさらに続ければ，$Q=Q^{(1)}Q^{(2)}\cdots Q^{(n-1)}$ として，(5.2.1)が得られる．

注意 5.2.1　$A^{(1)}\boldsymbol{y}=\lambda_2\boldsymbol{y}$ を満たす $\lambda_2, \boldsymbol{y}\,(\neq\boldsymbol{0})$ は $A^{(1)}$ の固有値，固有ベクトルであり，λ_2 は A の固有値でもあるが，$\boldsymbol{y}$ が A の固有値 λ_2 に対応する固有ベクトルに直接結びつけられるわけではない．実際，$\tilde{\boldsymbol{y}}=[z, \boldsymbol{y}^{\mathrm{T}}]^{\mathrm{T}}$ とおいて(5.2.2)に右から掛けると($\boldsymbol{a}'=\square$ として)

$$Q^{(1)\mathrm{H}}AQ^{(1)}\tilde{\boldsymbol{y}}=\begin{bmatrix}\lambda_1 z+\boldsymbol{a}'\boldsymbol{y}\\ \lambda_2\boldsymbol{y}\end{bmatrix} \tag{5.2.7}$$

となるので，

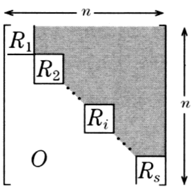
$$\lambda_1 z+\boldsymbol{a}'\boldsymbol{y}=\lambda_2 z \tag{5.2.8}$$

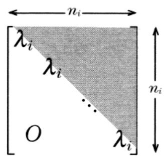
ならば $\tilde{\boldsymbol{y}}$ が A の(固有値 λ_2 に対応する)固有ベクトルとなる．$\lambda_1\neq\lambda_2$ なら $z=\boldsymbol{a}'\boldsymbol{y}/(\lambda_2-\lambda_1)$ により(5.2.8)を満たすように z を定めることができるが，$\lambda_1=\lambda_2$ のときにはそうは行かない．この辺の事情を以下で逐次解明して行く．

(5.2.1)の Schur 形を構成する際に，A に重複固有値があるときにはそれらを続けて処理することにすると，

$$\varphi_A(\lambda)=(\lambda_1-\lambda)^{n_1}\cdots(\lambda_s-\lambda)^{n_s}\qquad\left(\sum_{i=1}^{s}n_i=n\right) \tag{5.2.9}$$

であるときには

$R=$ 

，　$R_i=$ 

(5.2.10)

を得る．

いま，R を

$$R=\left[\begin{array}{c|c}R_1 & C\\ \hline O & \tilde{R}\end{array}\right] \tag{5.2.11}$$

と分割し，正則行列 T とその逆行列 T^{-1} を $n_1\times(n-n_1)$ 型の未知行列 X を含んだ

$$T=\left[\begin{array}{c|c} I_{n_1} & -X \\ \hline O & I_{n-n_1}\end{array}\right],\qquad T^{-1}=\left[\begin{array}{c|c} I_{n_1} & X \\ \hline O & I_{n-n_1}\end{array}\right] \tag{5.2.12}$$

という形に仮定して，

$$T^{-1}RT=\left[\begin{array}{c|c} R_1 & O \\ \hline O & \tilde{R}\end{array}\right]=R_1\dotplus\tilde{R} \tag{5.2.13}$$

(→(3.3.38))を満たすように X を定めることを試みよう．(5.2.13)は

$$R_1X-X\tilde{R}=C \tag{5.2.14}$$

であれば成り立つ．(5.2.14)は Sylvester の方程式(4.5.3 項)で，R_1 の固有値は R_1 の対角要素 λ_1 に等しく，$\tilde{R}$ の固有値は $\tilde{R}$ の対角要素 $\lambda_2,\cdots,\lambda_s$ に等しく，λ_1 と $\lambda_2,\cdots,\lambda_s$ とは異なるから，一意解 X をもつ．したがって，(5.2.13)を満たすように X，すなわち T を定めることができる．$QT=T'$ とおけば

$$T'^{-1}AT'=R_1\dotplus\tilde{R} \tag{5.2.15}$$

である．同様のことを $\tilde{R}$ に対して繰り返せば，結局，ある正則行列 S による相似変換で $S^{-1}AS=R_1\dotplus\cdots\dotplus R_s$ （R_i は(5.2.10)の形）に変換することができる．

注意 5.2.2 注意 5.1.6 によれば $\varphi_A(x)=\varphi_R(x)$ である．また，R^k は対角要素に $\lambda_1{}^k,\cdots,\lambda_n{}^k$ をもつから，$\varphi_{R^k}(x)=\varphi_{A^k}(x)$ により，A^k の固有値が $\lambda_1{}^k,\cdots,\lambda_n{}^k$ であることがわかる．さらに，多項式 $f(x)$ に対して，行列多項式 $f(A)$ の固有値が $f(\lambda_1),\cdots,f(\lambda_n)$ であることもわかる．

注意 5.2.3 A を $R_1\dotplus\cdots\dotplus R_s$ の形に相似変換する S はもはや，一般には，ユニタリ行列ではない．

表現行列が $R_1\dotplus\cdots\dotplus R_s$ の形となるように V_n の基底を選べば，V_n が各ブロック R_i に対応する(すなわち，値の異なる固有値に対応する)部分空間の直和 $V_n=U(\lambda_1)\oplus\cdots\oplus U(\lambda_s)$ に分解されることは明らかである．

注意 5.2.4 上の注意 5.2.3 の直和分解は，$R_1\dotplus\cdots\dotplus R_s$ の作り方にはよらず A のみによって定まるものである．いま，$S^{-1}(A-\lambda_iI)S=(R_1\dotplus\cdots\dotplus R_s)-\lambda_iI$ を作ると，第 i 対角ブロック $R_i-\lambda_iI_{n_i}$ は**狭義上三角**(strictly upper triangular)，すなわち

$$N_i\equiv R_i-\boldsymbol{\lambda}_iI_{n_i}=\begin{bmatrix} 0 & & \blacktriangle \\ & 0 & \\ & & \ddots \\ O & & & 0\end{bmatrix} \tag{5.2.16}$$

という形になるから，明らかに $N_i{}^{n_i}=O_{n_i,n_i}$ となる．そこで，$\boldsymbol{x}\in U(\lambda_i)$ なら $(A-\lambda_iI)^{n_i}\boldsymbol{x}=\boldsymbol{0}$ である．他方，$\lambda_j\neq\lambda_i$ に対しては $(A-\lambda_jI)^{n_i}\boldsymbol{x}\neq\boldsymbol{0}$ であることも容易に示される．このことは $U(\lambda_i)$ を「$(A-\lambda_iI)^{n_i}\boldsymbol{x}=\boldsymbol{0}$ を満たすベクトル $\boldsymbol{x}$ の集ま

り」として定義できることを示している．

$U(\lambda_i)$のことを，A の固有値 λ_i に対応する**固有空間**(eigenspace)と呼ぶことがある．

注意 5.2.5 何乗かして零行列となるような正方行列 A は**べき零**(nilpotent)であるといわれる．べき零行列の Schur 形が狭義上三角であることは明らかであろう．

注意 5.2.6 (5.2.16)の N_i は(ユニタリとは限らない)相似変換でさらに簡単な形にできる．記号の倹約のため N_i を単に N と書き，N($m\times m$ 型とする)を相似変換によって

$$S^{-1}NS = K_1 \dotplus \cdots \dotplus K_t, \quad K_j = \overbrace{\begin{bmatrix} 0 & 1 & & O \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ O & & & 0 \end{bmatrix}}^{m_j} \Big\updownarrow m_j, \tag{5.2.17}$$

$$m_1 \geqq m_2 \geqq \cdots \geqq m_t, \quad m_1 + m_2 + \cdots + m_t = m$$

という形にできることを示そう($m_j=1$ のときは $K_j=[0]$であるとする)．

m 次以下の N についてはこのような形に相似変換できたとして，$m+1$ 次の

$$N = \left[\begin{array}{c|c} 0 & \overbrace{\ast\ \cdots\ \ast}^{m} \\ \hline O & \begin{matrix} 0 & \ddots & \ast \\ & N^{(1)} & \\ O & \ddots & 0 \end{matrix} \end{array}\right] \Big\updownarrow m \tag{5.2.18}$$

について考える．$N^{(1)}$ の部分を(5.2.17)の形 $N^{(2)}$ に変換する m 次の正則行列を $T^{(1)}$ とすれば，

$$\left[\begin{array}{c|c} 1 & 0 \ \cdots \ 0 \\ \hline 0 & \\ \vdots & T^{(1)} \\ 0 & \end{array}\right]^{-1} N \left[\begin{array}{c|c} 1 & 0 \ \cdots \ 0 \\ \hline 0 & \\ \vdots & T^{(1)} \\ 0 & \end{array}\right] = \left[\begin{array}{c|c} 0 & a_1 \ a_2 \ \cdots \ a_m \\ \hline 0 & \begin{matrix} 0 & \ddots & \ast \\ & N^{(2)} & \\ O & \ddots & 0 \end{matrix} \\ \vdots & \\ 0 & \end{array}\right] = N', \tag{5.2.19}$$

$$N^{(2)} = \left[\begin{array}{c|c} \overset{m_1}{K_1} & \overset{m-m_1}{O} \\ \hline O & \begin{matrix} 0 & \ddots & \ast \\ & K^{(1)} & \\ O & \ddots & 0 \end{matrix} \end{array}\right] \begin{matrix} \updownarrow m_1 \\ \updownarrow m-m_1 \end{matrix}, \quad K_1 = \overbrace{\begin{bmatrix} 0 & 1 & & O \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ O & & & 0 \end{bmatrix}}^{m_1} \Big\updownarrow m_1 \tag{5.2.20}$$

という形になる．(ここで m_1 は $K^{(1)}$ の各対角ブロックの次数 m_j $(j\geqq 2)$以上である．したがって

$$(K^{(1)})^{m_1} = O \tag{5.2.21}$$

である．）さらに，$\boldsymbol{a}=[\overbrace{a_2, a_3, \cdots, a_{m_1}}^{m_1-1}, \overbrace{0, \cdots\cdots\cdots, 0}^{m-m_1+1}]$ とおいて，$m+1$ 次の正則行列

$$T^{(2)}=\left[\begin{array}{c|c} 1 & \boldsymbol{a} \\ \hline 0 & \\ \vdots & I_m \\ 0 & \end{array}\right], \qquad (T^{(2)})^{-1}=\left[\begin{array}{c|c} 1 & -\boldsymbol{a} \\ \hline 0 & \\ \vdots & I_m \\ 0 & \end{array}\right] \tag{5.2.22}$$

により N' を相似変換すれば

$$N^{[0]}=(T^{(2)})^{-1}N'T^{(2)}=\left[\begin{array}{c|c|c} 0 & a_1 0\cdots 0 & a_{m_1+1}\cdots a_m \\ \hline 0 & & \\ \vdots & K_1 & O \\ 0 & & \\ \hline 0 & & \\ \vdots & O & K^{(1)} \\ 0 & & \end{array}\right] \begin{array}{l} \\ \updownarrow m_1 \\ \\ \updownarrow m-m_1 \end{array} \tag{5.2.23}$$

（列幅：$\overleftrightarrow{m_1}$，$\overleftrightarrow{m-m_1}$）

という形になる．

〈$a_1 \neq 0$ のとき〉 $N^{[0]}$ の行 1 を a_1 で割り，列 1 を a_1 倍すること（相似変換である；→ 2.3.2 項の $T(i_0 ; t)$ の形の基本行列と関係(2.3.10)）により $N^{[0]}$ は

$$N^{[1]}=\left[\begin{array}{c|c} & \boldsymbol{b}^{[1]} \\ \cline{2-2} K_1' & O \\ \hline O & K^{(1)} \end{array}\right] \begin{array}{l} \updownarrow m_1+1 \\ \updownarrow m-m_1 \end{array}, \qquad K_1'=\begin{bmatrix} 0 & 1 & & O \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ O & & & 0 \end{bmatrix} \updownarrow m_1+1 \tag{5.2.24}$$

（$N^{[1]}$ の列幅：$\overleftrightarrow{m_1+1}$，$\overleftrightarrow{m-m_1}$；$K_1'$ の列幅：$\overleftrightarrow{m_1+1}$）

$(\boldsymbol{b}^{[1]}=[b^{[1]}_{m_1+1}, \cdots, b^{[1]}_m]$；$b^{[1]}_j=a_j/a_1$，$j=m_1+1, \cdots, m)$ となる．$N^{[1]}$ の列 2 を $b^{[1]}_{m_1+1}$ 倍，…，$b^{[1]}_m$ 倍したものを，それぞれ，列 m_1+2，…，列 $m+1$ から引き，行 m_1+2，…，行 $m+1$ を，それぞれ，$b^{[1]}_{m_1+1}$ 倍，…，$b^{[1]}_m$ 倍したものを行 2 に加えること（これも相似変換；→ 2.3.2 項，特に $W(i_0, j_0 ; t)$ の形の基本行列と関係(2.3.10)）により，$N^{[1]}$ は

$$N^{[2]}=\left[\begin{array}{c|c} & 0\cdots\cdots 0 \\ \cline{2-2} & 0 \;\; \boldsymbol{b}^{[2]} \\ \cline{2-2} K_1' & O \\ \hline O & K^{(1)} \end{array}\right], \qquad \boldsymbol{b}^{[2]}=[b^{[2]}_{m_1+2}, \cdots, b^{[2]}_m] \tag{5.2.25}$$

に変換される．ここで，$K^{(1)}$ の対角線の一つ上の要素が 1 であるか 0 であるかによって，$b^{[2]}_j=b^{[1]}_{j-1}$ あるいは $b^{[2]}_j=0$ である．まとめて書けば

$$[0, \boldsymbol{b}^{[2]}]=\boldsymbol{b}^{[1]}\cdot K^{(1)} \tag{5.2.26}$$

と書ける．さらに，$j=m_1+2, \cdots, m$ について，$N^{[2]}$ の列 3 を $b^{[2]}_j$ 倍したものを列 j

$+1$ から引き，行 $j+1$ を $b_j^{[2]}$ 倍したものを行 3 に加える；……；ということを続ける．すると，$m_1 \geqq m-m_1$ $(2m_1 \geqq m)$ ならば

$$N^{[m-m_1+1]} = \begin{array}{c} \scriptstyle \leftarrow m_1+1 \rightarrow \leftarrow m-m_1 \rightarrow \\ \left[\begin{array}{c|c} K_1' & O \\ \hline O & K^{(1)} \end{array}\right] \end{array} \begin{array}{l} \scriptstyle \updownarrow m_1+1 \\ \scriptstyle \updownarrow m-m_1 \end{array} \tag{5.2.27}$$

が所望の形(5.2.17)となっているし，$m_1 < m-m_1$ $(2m_1 < m)$ ならば

$$N^{[m_1+1]} = \begin{array}{c} \scriptstyle \leftarrow m_1+1 \rightarrow \leftarrow m_1 \rightarrow \leftarrow m-2m_1 \rightarrow \\ \left[\begin{array}{c|cc} K_1' & & O \\ & & \boldsymbol{b}^{[m_1+1]} \\ \hline O & & K^{(1)} \end{array}\right] \end{array} \begin{array}{l} \scriptstyle \updownarrow m_1+1 \\ \scriptstyle \updownarrow m-m_1 \end{array} \tag{5.2.28}$$

となるが，

$$[\overbrace{0,\cdots,0}^{m_1},\ \overbrace{\boldsymbol{b}^{[m_1+1]}}^{m-2m_1}] = \boldsymbol{b}^{[1]} \cdot (K^{(1)})^{m_1} \tag{5.2.29}$$

であるから(5.2.21)により $\boldsymbol{b}^{[m_1+1]} = \boldsymbol{0}$ で，やはり所望の形(5.2.17)を得る．

〈$a_1=0$ のとき〉 (5.2.23)の列 $1, 2, \cdots, m_1, m_1+1$ を $2, \cdots, m_1, m_1+1, 1$ の順に並べかえ，同時に行も同様に並べかえること(これも相似変換；→ 2.3.2 項)によって，

$$N^{[1]} = \begin{array}{c} \scriptstyle \leftarrow m_1 \rightarrow \leftarrow m-m_1+1 \rightarrow \\ \left[\begin{array}{c|cc} K_1 & & O \\ \hline & 0 & a_{m_1+1} \cdots a_m \\ O & 0 & \\ & \vdots & K^{(1)} \\ & 0 & \end{array}\right] \end{array} \begin{array}{l} \scriptstyle \updownarrow m_1 \\ \scriptstyle \updownarrow m-m_1+1 \end{array} \tag{5.2.30}$$

となるが，(5.2.30)の影をつけた部分は m 次以下の上三角行列(対角要素は 0)であるから，帰納法の仮定により，再び(5.2.17)の形に相似変換できる．その結果生じる対角ブロックの中に次数が m_1 を超えるものがあれば，ブロックの順序を入れ換えることによって，$N^{[1]}$ 全体を所望の形(5.2.17)にすることができる．

注意 5.2.7 注意 5.2.6 により，任意の n 次正方行列 A は次の形の行列 J に相似変換できることがわかる．ここで，$\varphi_A(\lambda) = \prod_{i=1}^{s} (\lambda_i - \lambda)^{n_i}$ $\left(\sum_{i=1}^{s} n_i = n\right)$ とする．

$$J = J_1 \dot{+} \cdots \dot{+} J_s \qquad (J_i \text{ は } n_i \text{ 次正方}), \tag{5.2.31a}$$

$$J_i = J_{i1} \dot{+} \cdots \dot{+} J_{is_i}$$

$$\left(J_{ij} \text{ は } n_{ij} \text{ 次正方；} \sum_{j=1}^{s_i} n_{ij} = n_i \text{；} n_{i1} \geqq n_{i2} \geqq \cdots \geqq n_{is_i}\right), \tag{5.2.31b}$$

$$J_{ij} = \begin{bmatrix} \lambda_i & 1 & & & O \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \ddots & 1 \\ O & & & & \lambda_i \end{bmatrix}. \tag{5.2.31c}$$

ここで，(5.2.31a)の各対角ブロックに対応する基底ベクトルの張る部分空間 $U(\lambda_i)$が一意に定まることは注意 5.2.3 で述べた通りであるが，(5.2.31b)の場合にはこのような一意性はない(→ 演習問題 5.3)．しかし 6.8.1，6.8.2 項で示すように，A に対して s，λ_i，n_i，$U(\lambda_i)$ $(i=1,\cdots,s)$，s_i，n_{ij} $(j=1,\cdots,s_i\,;\,i=1,\cdots,s)$は一意に定まる．

注意 5.2.8 上の注意 5.2.7 で述べた事実は，n 次正方行列 A の固有値 λ_i に対応する固有空間 $U_i=U(\lambda_i)$に属する固有ベクトルなどの構造に関する詳しい情報も与えてくれる$(i=1,\cdots,s)$．すなわち，相異なる固有値 $\lambda_1,\cdots,\lambda_s$(多重度はそれぞれ $n_1,\cdots,n_s$)の各々に対応する固有空間 U_i (一意に定まる；(5.2.31a)の J_i に対応する)の中で，各 $J_{ij}(j=1,\cdots,s_i)$に対応する部分空間 U_{ij} (これは注意 5.2.7 で述べたように一意には定まらないが)において(5.2.31b)，(5.2.31c)の形の A の表現を実現する基底の基底ベクトルを $\boldsymbol{e}_1^{(i,j)},\cdots,\boldsymbol{e}_{n_{ij}}^{(i,j)}$ と記すと，(5.2.31)は

$$A\boldsymbol{e}_1^{(i,j)}=\lambda_i\boldsymbol{e}_1^{(i,j)} \tag{5.2.32}$$

を表しているので，$\boldsymbol{e}_1^{(i,j)}$ は固有値 λ_i に対応する“一つの”固有ベクトルであるが，他の基底ベクトルは

$$A\boldsymbol{e}_j^{(i,j)}=\lambda_i\boldsymbol{e}_j^{(i,j)}+\boldsymbol{e}_{j-1}^{(i,j)} \qquad (j=2,\cdots,n_{ij}) \tag{5.2.33}$$

を満たしている．このように，n_i 重の固有値 λ_i に対応して，s_i 個の固有ベクトル $\boldsymbol{e}_1^{(i,j)}(j=1,\cdots,s_i)$が存在し，それ以外に，各 $\boldsymbol{e}_1^{(i,j)}$ と関係(5.2.33)で結ばれるベクトル $\boldsymbol{e}_2^{(i,j)},\cdots,\boldsymbol{e}_{n_{ij}}^{(i,j)}$(これらを**一般化固有ベクトル**(generalized eigenvectors)と呼ぶ)が存在する．固有値 λ_i に対応する独立な固有ベクトルの数 $s_i=\dim\mathrm{Ker}(A-\lambda_i I)$ を固有値 λ_i の**幾何学的多重度**(geometric multiplicity)と呼び，**代数的多重度**(algebraic multiplicity) n_i と区別する流儀がある．また，各 $\boldsymbol{e}_1^{(i,j)}$ に属する一般化固有ベクトルの個数 n_{ij} を，部分空間 U_{ij} の**高さ**(height)，$\max_j n_{ij}$ を固有値 λ_i(あるいは固有空間 U_i)の高さ，などと呼ぶこともある．

5.2.2 ユニタリ相似(ユニタリ合同)変換による行列の対角化

Q がユニタリ行列ならば $Q^{\mathrm{H}}=Q^{-1}$ であるから，ユニタリ相似変換 $Q^{-1}AQ$ とユニタリ合同変換 $Q^{\mathrm{H}}AQ$ とは一致する．また，A が正規行列(2.14.3 項)であれば，それにユニタリ相似な行列 $Q^{\mathrm{H}}AQ$ も正規行列である．実際，

$$(Q^{\mathrm{H}}AQ)(Q^{\mathrm{H}}AQ)^{\mathrm{H}}=Q^{\mathrm{H}}AQQ^{\mathrm{H}}A^{\mathrm{H}}Q=Q^{\mathrm{H}}AA^{\mathrm{H}}Q$$
$$=Q^{\mathrm{H}}A^{\mathrm{H}}AQ=(Q^{\mathrm{H}}AQ)^{\mathrm{H}}(Q^{\mathrm{H}}AQ). \qquad (5.2.34)$$

したがって，A が正規なら，その Schur 形 R も正規である．ところが，上三角行列 $R=[r_{ij}]$ が正規であるとすると，

$$\begin{bmatrix} r_{11} & & r_{ij} \\ & r_{22} & \\ & O & \ddots \\ & & & r_{nn} \end{bmatrix}\begin{bmatrix} \bar{r}_{11} & & O \\ & \bar{r}_{22} & \\ & \bar{r}_{ji} & \ddots \\ & & & \bar{r}_{nn} \end{bmatrix}=\begin{bmatrix} \bar{r}_{11} & & O \\ & \bar{r}_{22} & \\ & \bar{r}_{ji} & \ddots \\ & & & \bar{r}_{nn} \end{bmatrix}\begin{bmatrix} r_{11} & & r_{ij} \\ & r_{22} & \\ & O & \ddots \\ & & & r_{nn} \end{bmatrix} \qquad (5.2.35)$$

の (i,i) 要素を比較することにより，

$$\sum_{j=i+1}^{n}|r_{ij}|^2=\sum_{k=1}^{i-1}|r_{ki}|^2 \qquad (5.2.36)$$

であるから，$i=1,2,3,\cdots$ の順に $\sum_{j=i+1}^{n}|r_{ij}|^2=0$ でなければならないこと，したがって $r_{ij}=0\,(j=i+1,\cdots,n)$ であることがわかる．すなわち，R は対角行列でなければならない．

このようにして，

「正規行列 A はユニタリ相似変換(＝ユニタリ合同変換)によって，固有値を対角要素としてもつ対角行列にすることができる」

$$Q^{\mathrm{H}}AQ=\mathrm{diag}[\lambda_1,\lambda_2,\cdots,\lambda_n] \qquad (5.2.37)$$

という重要な定理が得られる(対角行列は正規行列であるから，この逆「ユニタリ相似変換で対角化される行列は正規行列である」は明らか)，(5.2.37) は

$$AQ=Q\cdot\mathrm{diag}[\lambda_1,\lambda_2,\cdots,\lambda_n] \qquad (5.2.38)$$

と書き直せば，

「Q の列 j が固有値 λ_i に対応する固有ベクトルである」

ことを示しているので，

「正規行列 A は n 本の直交系をなす固有ベクトル(もちろん独立)をもつ」

こともわかる．さらに，$\|\mathrm{diag}[\lambda_1,\cdots,\lambda_n]\|_2=\max|\lambda_i|$ であり，ユニタリ変換でベクトルの 2 乗ノルムは変化しないから「正規行列 A のスペクトル半径 $\rho(A)$ は A の 2 乗ノルム $\|A\|_2$ に等しい」ことになる．

例 5.2.1(正規行列の例)　(i)　A が Hermite 行列のとき，$A^{\mathrm{H}}=A$．そこで，

(5.2.37)の両辺の Hermite 共役をとれば $Q^{H}A^{H}Q=Q^{H}AQ$ であるから，$\bar{\lambda}_i=\lambda_i$，すなわち Hermite 行列の固有値は実数である．左固有ベクトルと右固有ベクトルが一致することも $A=A^{H}$ だから明らか．

正規化された固有ベクトルを基底に取れば，A を係数行列とする Hermite 形式は

$$\boldsymbol{x}^{H}A\boldsymbol{x}=\lambda_1|x_1|^2+\cdots+\lambda_n|x_n|^2 \tag{5.2.39}$$

と書かれるから，A の Sylvester の符号指数は，(正の固有値の個数，負の固有値の個数，零固有値の個数)とも表せることになる(→ 注意 3.5.9)．

(ii) A が歪 Hermite 行列のとき，$A^{H}=-A$．$\mathrm{i}A$ が Hermite 行列であるから，歪 Hermite 行列の固有値は純虚数である．

(iii) A がユニタリ行列のとき，$A^{H}=A^{-1}$．$(Q^{H}AQ)^{H}(Q^{H}AQ)=Q^{H}A^{H}AQ=Q^{H}Q=I_n$．一方，$\mathrm{diag}[\bar{\lambda}_1,\cdots,\bar{\lambda}_n]\cdot\mathrm{diag}[\lambda_1,\cdots,\lambda_n]=\mathrm{diag}[|\lambda_1|^2,\cdots,|\lambda_n|^2]$．したがって $|\lambda_i|^2=1$，$|\lambda_i|=1$．すなわち，ユニタリ行列の固有値の絶対値は 1 である($\lambda_i=\mathrm{e}^{\mathrm{i}\sigma_i}$，$\sigma_i$：実数，という形に表せる)． □

5.2.3 Gershgorin の定理と一般化

n 次正方行列 A の固有値の存在範囲をきわめて簡単な形で表す方法の一つに，次の有名な定理がある．

Gershgorin(ゲルシュゴリン)の定理(弱い形) n 次正方行列 $A=[a_{ij}]\,(i,j=1,\cdots,n)$ の任意の固有値 λ は，

$$C_i=\{z\in\mathbb{C}\mid |z-a_{ii}|\leqq r_i\},\qquad r_i=\sum_{j\neq i}|a_{ij}| \tag{5.2.40}$$

で定義される，複素平面上の n 個の円 $C_i\,(i=1,\cdots,n)$ (それぞれ，中心 a_{ii} 半径 r_i)の合併集合 $\bigcup_{i=1}^{n}C_i$ の中にある．

証明 A の一つの固有値 λ とそれに対応する一つの固有ベクトルを $\boldsymbol{x}=[x_j]$ とする：

$$A\boldsymbol{x}=\lambda\boldsymbol{x}\,;\qquad \sum_{j=1}^{n}a_{ij}x_j=\lambda x_i\quad(i=1,\cdots,n). \tag{5.2.41}$$

$\boldsymbol{x}\neq\boldsymbol{0}$ であるから，$|x_{i_0}|=\max_j|x_j|\neq 0$ とおくと，(5.2.41)の $i=i_0$ に対する式より

$$(\lambda-a_{i_0i_0})\,x_{i_0}=\sum_{j\neq i_0}a_{i_0j}x_j, \tag{5.2.42}$$

したがって

$$|\lambda - a_{i_0 i_0}| \leqq \sum_{j \neq i_0} |a_{i_0 j}| \frac{|x_j|}{|x_{i_0}|} \leqq \sum_{j \neq i_0} |a_{i_0 j}| = r_{i_0} \tag{5.2.43}$$

を得る．これは $\lambda \in C_{i_0} \subseteqq \bigcup_{i=1}^{n} C_i$ を意味する． ■

特性多項式 $\varphi_A(\lambda)$ の零点は，その多項式の係数の，したがって A の要素 a_{ij} の“連続関数”であることに着目すると，上の弱い形からより強い次の定理が導き出せる．

Gershgorin の定理(強い形)　$A=[a_{ij}]$ の n 個の固有値 $\lambda_1,\cdots,\lambda_n$ (多重固有値は多重度だけ繰り返して数える)は，(5.2.40)で定義される n 個の円 $C_1,\cdots,C_n$ の合併集合 $\bigcup_{i=1}^{n} C_i$ の各連結成分の中に，それを構成する円の個数に等しい個数ずつ含まれる．

略 証　$A=D+B$ ($D=\mathrm{diag}[a_{11},\cdots,a_{nn}]$；$B=[b_{ij}], b_{ii}=0, b_{ij}=a_{ij} (i \neq j)$) とし，$t$ をパラメタとする行列の族 $A(t)=D+tB$ ($t \in [0,1]$) を考える．$A(0)=D, A(1)=A$ である．この強い形の定理が $A(0)$ に対して成立することは明らかである．$A(t)$ に対応する円 $C_i(t)$ は，t と共に連続的に変化し，λ_i も t の連続関数であるから，弱い形の定理が成り立てば強い形の定理も成り立つ． ■

注意 5.2.9　(5.2.40)の円 $C_i (i=1,\cdots,n)$ を **Gershgorin の円** (Gershgorin circles) という．上の定理は「$\bigcup_{i=1}^{n} C_i$ の各連結成分の中にその連結成分を構成する円の個数に等しい数の固有値が含まれている」と述べているのであって，「各円 C_i の中に固有値が1個ずつある」といっているわけではない(→ 演習問題 5.4)．

n 次正方行列 $A=[a_{ij}]$ の n 個の固有値の中から m 個 $\lambda_{i_1},\cdots,\lambda_{i_m}$ ($i_1<i_2<\cdots<i_m$) を選んでその平均値を

$$\mu(i_1,\cdots,i_m) = \frac{1}{m} \sum_{k=1}^{m} \lambda_{i_k} \tag{5.2.44}$$

とする．行列 A の行 $j_1,\cdots,j_m$ と列 $j_1,\cdots,j_m$ ($j_1<\cdots<j_m$) からなる部分正方行列 $A_{j_1\cdots j_m}^{j_1\cdots j_m}$ のトレースの $1/m$ を $\alpha(j_1,\cdots,j_m)$，また

$$r(j_1,\cdots,j_m) = \frac{1}{m} \sum_{\substack{i \in \{j_1,\cdots,j_m\} \\ j \notin \{j_1,\cdots,j_m\}}} |a_{ij}| \tag{5.2.45}$$

として，**一般化された Gershgorin の円** (generalized Gershgorin circles) を

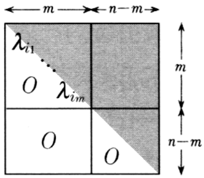
$$C(j_1,\cdots,j_m) = \{z \in \mathbb{C} \mid |z - \alpha(j_1,\cdots,j_m)| \leqq r(j_1,\cdots,j_m)\} \tag{5.2.46}$$

で定義する．すると次の定理が成り立つ．

一般化された Gershgorin の定理 $\binom{n}{m}$ 個の平均固有値 $\mu(i_1,\cdots,i_m)$ は，$\binom{n}{m}$ 個の $C(j_1,\cdots,j_m)$ の合併集合の中にある．特に合併集合の一つの連結成分が l 個の円からなるとき，その連結成分の中には l 個の平均固有値が含まれる．

証明 A の Schur 形を，$\lambda_{i_1},\cdots,\lambda_{i_m}$ が左上から順に並ぶように作り，

$$R=Q^{\mathrm{H}}AQ=\begin{bmatrix} R_{11} & R_{12} \\ O & R_{22} \end{bmatrix}=$$



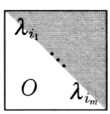
$$,\quad Q=\begin{bmatrix} S & T \end{bmatrix} \qquad (5.2.47)$$

とする．S の m 次の小行列式の中で絶対値が最大のものを $|S|_{1\cdots m}^{j_1\cdots j_m}$ $(j_1<\cdots<j_m)$ とする．簡単のため，A の行と列の順序を入れ換え，Q の行の順序を入れ換えて，$j_1,\cdots,j_m$ が先頭にくるようにし，

$$P^{\mathrm{T}}AP=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix},\quad P^{\mathrm{T}}Q=\begin{bmatrix} S_1 & T_1 \\ S_2 & T_2 \end{bmatrix} \qquad (5.2.48)$$

(P は置換行列)と書けば，(5.2.47)に等価な関係式 $(P^{\mathrm{T}}AP)(P^{\mathrm{T}}Q)=(P^{\mathrm{T}}Q)R$ の一部として

$$S_1R_{11}=S_1$$



$$=A_{11}S_1+A_{12}S_2, \qquad (5.2.49)$$

すなわち

$$S_1R_{11}S_1^{-1}=A_{11}+A_{12}S_2S_1^{-1} \qquad (5.2.50)$$

が得られる．この両辺のトレースをとると，

$\mathrm{tr}(S_1R_{11}S_1^{-1})=\mathrm{tr}\,R_{11}=m\mu(i_1,\cdots,i_m)=m\alpha(j_1,\cdots,j_m)+\mathrm{tr}(A_{12}S_2S_1^{-1})$，したがって，

$$|\mu(i_1,\cdots,i_m)-\alpha(j_1,\cdots,j_m)|=\frac{1}{m}|\mathrm{tr}(A_{12}S_2S_1^{-1})| \tag{5.2.51}$$

を得る．ところで，$Y=S_2S_1^{-1}$ は $(n-m)\times m$ 型行列であるが，その第 j 行ベクトル $\boldsymbol{y}_{j\cdot}$ は $\boldsymbol{y}_{j\cdot}S_1=\boldsymbol{s}_{j\cdot}$ という形の線形方程式系の解とみなせる（$\boldsymbol{s}_{j\cdot}$ は S_2 の第 j 行）ので，Cramer の公式(4.2 節)により

$$y_{ji}=\det[S_1\text{の第}\ i\ \text{行を}\ \boldsymbol{s}_{j\cdot}\ \text{でおきかえた行列}]/\det S_1$$
$$(j=1,\cdots,n-m\ ;\ i=1,\cdots,m) \tag{5.2.52}$$

と表せる．ここで，S_1 の定義から

$$|y_{ji}|\leqq 1 \tag{5.2.53}$$

であることがわかる．$\mathrm{tr}(A_{12}S_2S_1^{-1})=\mathrm{tr}(A_{12}Y)$ は，A_{12} の各要素にそれぞれ対応する Y の要素を掛けて和を取ったものに等しい(→(2.7.3))から，

$$|\mathrm{tr}(A_{12}S_2S_1^{-1})|\leqq A_{12}\text{の要素の絶対値の和} \tag{5.2.54}$$

となる．(5.2.51)，(5.2.54)，(5.2.45)より定理が証明されたことになる．■

5.2.4 Hermite 行列の固有値に関する最大・最小定理，Rayleigh 商

ここでは n 次 Hermite 行列だけを扱う．5.2.2 項で示したように，A には n 個の実数の固有値(それらを $\lambda_1\geqq\lambda_2\geqq\cdots\geqq\lambda_n$ とする)と対応する固有ベクトル(n 本を直交系をなすように選ぶことができる)が存在する．$\boldsymbol{x}_i\,(\neq\boldsymbol{0})$ が固有値 λ_i に対応する固有ベクトルであれば，$A\boldsymbol{x}_i=\lambda_i\boldsymbol{x}_i$ であるから，$\boldsymbol{x}_i^{\mathrm{H}}A\boldsymbol{x}_i=\lambda_i(\boldsymbol{x}_i^{\mathrm{H}}\cdot\boldsymbol{x}_i)$，$\lambda_i=(\boldsymbol{x}_i^{\mathrm{H}}A\boldsymbol{x}_i)/(\boldsymbol{x}_i^{\mathrm{H}}\boldsymbol{x}_i)$ である．そこで，固有ベクトルとは限らないベクトル $\boldsymbol{x}\,(\neq\boldsymbol{0})$ に対して **Rayleigh**(レイリー)**商**(Rayleigh quotient)を

$$R_A(\boldsymbol{x})=\frac{\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}} \tag{5.2.55}$$

で定義する．$\boldsymbol{x}$ が固有ベクトル $\boldsymbol{x}_i$ からわずかにずれている，すなわち，$\|\boldsymbol{\varepsilon}\|$ が小さい任意のベクトル $\boldsymbol{\varepsilon}$ に対して

$$\boldsymbol{x}=\boldsymbol{x}_i+\boldsymbol{\varepsilon} \tag{5.2.56}$$

とおくと，

$$\begin{aligned}R_A(\boldsymbol{x})&=\frac{(\boldsymbol{x}_i+\boldsymbol{\varepsilon})^{\mathrm{H}}A(\boldsymbol{x}_i+\boldsymbol{\varepsilon})}{(\boldsymbol{x}_i+\boldsymbol{\varepsilon})^{\mathrm{H}}(\boldsymbol{x}_i+\boldsymbol{\varepsilon})}\\&=\frac{\lambda_i(\|\boldsymbol{x}_i\|_2^2+2\mathrm{Re}(\boldsymbol{\varepsilon}^{\mathrm{H}}\boldsymbol{x}_i))+\boldsymbol{\varepsilon}^{\mathrm{H}}A\boldsymbol{\varepsilon}}{(\|\boldsymbol{x}_i\|_2^2+2\mathrm{Re}(\boldsymbol{\varepsilon}^{\mathrm{H}}\boldsymbol{x}_i))+\|\boldsymbol{\varepsilon}\|_2^2}\\&=\lambda_i+\mathrm{O}(\|\boldsymbol{\varepsilon}\|^2)\end{aligned} \tag{5.2.57}$$

となる．このことは，「$R_A(\boldsymbol{x})$ は $\boldsymbol{x}=\boldsymbol{x}_i$ において停留値 λ_i を取る」ことを示している．この性質があるため，Rayleigh 商は，固有値の近似計算，近似度の評価，A の摂動に対する固有値，固有ベクトルの変動の見積りなどに有用な道具となる．

次の定理は古くから応用上重要な役割を果たしつづけている．

Hermite 行列の固有値に関する最大・最小定理 $A: V_n \to V_n$ が Hermite で，固有値 $\lambda_1 \geqq \cdots \geqq \lambda_n$ をもつとする．V_n の任意の k 次元部分空間を U_k と記す($k=1,\cdots,n$)．このとき次の(i)，(ii)が成り立つ．

(i) $$\lambda_k = \max_{U_k} \min\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x} \in U_k\}. \tag{5.2.58}$$

(ii) $$\lambda_{n-k+1} = \min_{U_k} \max\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x} \in U_k\}. \tag{5.2.59}$$

注意 5.2.10 $k=1$ のときは $R_A(\boldsymbol{x})$ $(\boldsymbol{x} \in U_1)$ は U_1 のみで定まるから

$$\lambda_1 = \max_{\boldsymbol{x} \neq \boldsymbol{0}} R_A(\boldsymbol{x}), \qquad \lambda_n = \min_{\boldsymbol{x} \neq \boldsymbol{0}} R_A(\boldsymbol{x}), \tag{5.2.60}$$

すなわち，Rayleigh 商の最大値，最小値は A の最大固有値，最小固有値に一致する．$k=n$ のときを考えても同じ結果になる．

証明 A をユニタリ相似変換により対角化し(→ 5.2.2 項)対角要素 λ_i が大きさの順に並ぶようにする：

$$Q^{\mathrm{H}}AQ = \mathrm{diag}[\lambda_1, \cdots, \lambda_n], \qquad \lambda_1 \geqq \cdots \geqq \lambda_n. \tag{5.2.61}$$

$\boldsymbol{x}=Q\boldsymbol{y}$，$\boldsymbol{y}=Q^{\mathrm{H}}\boldsymbol{x}$ とおくと，

$$R_A(\boldsymbol{x}) = \frac{\sum_{i=1}^{n} \lambda_i |y_i|^2}{\sum_{i=1}^{n} |y_i|^2} \tag{5.2.62}$$

である．さらに，V_n の部分空間 $V_1 \subset \cdots \subset V_{n-1} \subset V_n$ を

$$V_k = \{\boldsymbol{x} \mid y_{k+1} = \cdots = y_n = 0\} \tag{5.2.63}$$

により，$V_n \supset \bar{V}_{n-1} \supset \cdots \supset \bar{V}_1$ を

$$\bar{V}_k = \{\boldsymbol{x} \mid y_1 = \cdots = y_{n-k} = 0\} \tag{5.2.64}$$

で定義する．(A を対角行列に表現する，すなわち A の固有ベクトルからなる V_n の正規直交基底を$\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_n\}$とすると

$$V_k = L(\boldsymbol{e}_1, \cdots, \boldsymbol{e}_k), \tag{5.2.65}$$

$$\bar{V}_k = L(\boldsymbol{e}_{n-k+1}, \cdots, \boldsymbol{e}_n) \tag{5.2.66}$$

である.) $\boldsymbol{x} \in V_k$ なら

$$R_A(\boldsymbol{x}) = (\lambda_1|y_1|^2 + \cdots + \lambda_k|y_k|^2) / (|y_1|^2 + \cdots + |y_k|^2) \tag{5.2.67}$$

であるから,

$$\boldsymbol{x} \in V_k \quad \text{なら} \quad \lambda_1 \geqq R_A(\boldsymbol{x}) \geqq \lambda_k. \tag{5.2.68}$$

同様に,

$$\boldsymbol{x} \in \bar{V}_k \quad \text{なら} \quad \lambda_{n-k+1} \geqq R_A(\boldsymbol{x}) \geqq \lambda_n. \tag{5.2.69}$$

(これらの不等号は, それぞれ $\boldsymbol{x} = \boldsymbol{e}_1, \boldsymbol{e}_k, \boldsymbol{e}_{n-k+1}, \boldsymbol{e}_n$ に対して等号となる.)

任意の k 次元部分空間 U_k については, 最小値を取る範囲が狭くなれば最小値は大きくなる(小さくはならない)から,

$$\min\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x} \in U_k\} \leqq \min\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x} \in U_k \cap \bar{V}_{n-k+1}\}. \tag{5.2.70}$$

一方, $\dim(U_k \cap \bar{V}_{n-k+1}) + \dim(U_k + \bar{V}_{n-k+1}) = \dim U_k + \dim \bar{V}_{n-k+1} = k + (n-k+1) = n+1$ (→(3.2.3)), $\dim(U_k + \bar{V}_{n-k+1}) \leqq \dim V_n = n$ ($U_k + \bar{V}_{n-k+1} \subseteqq V_n$ と注意 3.2.2)より, $\dim(U_k \cap \bar{V}_{n-k+1}) \geqq 1$ であり, したがって $U_k \cap \bar{V}_{n-k+1} \neq \{\boldsymbol{0}\}$, すなわち $U_k \cap \bar{V}_{n-k+1}$ はある非零ベクトル $\boldsymbol{x}$ を含む. $\boldsymbol{x} \in \bar{V}_{n-k+1}$ でもあるから, (5.2.69)より, 任意の U_k に対して

$$\min\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x} \in U_k \cap \bar{V}_{n-k+1}\} \leqq \lambda_k. \tag{5.2.71}$$

ゆえに, (5.2.71)と(5.2.70)より

$$\max_{U_k} \min\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x} \in U_k\} \leqq \lambda_k. \tag{5.2.72}$$

特に $U_k = V_k$ に選べば, (5.2.68)より $\min\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x} \in V_k\} = \lambda_k$ であるから, (5.2.72)は等号で成立する. これで(i)の証明が完了した. (ii)の証明も同様. ∎

Hermite 行列 A を係数行列とする V_n における二次形式は, V_n の部分空間 W_m $(m \leqq n)$ における二次形式を定める. すなわち, W_m の正規直交基底の一つを $\{\boldsymbol{c}_1, \cdots, \boldsymbol{c}_m\}$ とし $n \times m$ 型の行列 $C = [\boldsymbol{c}_1, \cdots, \boldsymbol{c}_m]$ を作れば, $\boldsymbol{x} \in W_m$ は $\boldsymbol{x} = C\boldsymbol{y}$ と表される($\boldsymbol{y}$ は任意の m 次元縦ベクトル)ので, W_m においては

$$\boldsymbol{x}^{\mathrm{H}} A \boldsymbol{x} = \boldsymbol{y}^{\mathrm{H}} C^{\mathrm{H}} A C \boldsymbol{y} = \boldsymbol{y}^{\mathrm{H}} B \boldsymbol{y}, \qquad B = C^{\mathrm{H}} A C \tag{5.2.73}$$

で二次形式 $\boldsymbol{y}^{\mathrm{H}} B \boldsymbol{y}$ が定められる. ここで, $\{\boldsymbol{c}_1, \cdots, \boldsymbol{c}_m\}$ の正規直交性より $C^{\mathrm{H}} C = I_m$ であるから

$$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=\boldsymbol{y}^{\mathrm{H}}C^{\mathrm{H}}C\boldsymbol{y}=\boldsymbol{y}^{\mathrm{H}}\boldsymbol{y} \tag{5.2.74}$$

であることに注意する．このとき A の固有値 $\lambda_1 \geqq \cdots \geqq \lambda_n$ と B の固有値 $\mu_1 \geqq \cdots \geqq \mu_m$ の間には

$$\lambda_k \geqq \mu_k \geqq \lambda_{n-m+k} \qquad (k=1, \cdots, m), \tag{5.2.75}$$

あるいは，同じことであるが，

$$\lambda_{m-k} \geqq \mu_{m-k} \geqq \lambda_{n-k} \qquad (k=0, \cdots, m-1) \tag{5.2.75'}$$

という関係が成り立つ．

証明　この種の関係は Rayleigh 商を用いると見やすい．W_m における B の第 k 固有値 μ_k を最大・最小定理の(i)の形

$$\mu_k=\max_{U_k} \min\{R_B(\boldsymbol{y}) \mid \boldsymbol{y} \neq \boldsymbol{0}, \boldsymbol{y} \in U_k\}=\min\{R_B(\boldsymbol{y}) \mid \boldsymbol{y} \neq \boldsymbol{0}, \boldsymbol{y} \in \hat{U}_k\} \tag{5.2.76}$$

に表しておく(U_k は W_m の部分空間，$\hat{U}_k$ は中辺で max を達成するような部分空間)．$\hat{U}_k$ を V_n の部分空間とみなせば，A の第 k 固有値 λ_k は，U_k を V_n の部分空間として，

$$\begin{aligned}\lambda_k&=\max_{U_k} \min\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x} \in U_k\}\\ &\geqq \min\{R_A(\boldsymbol{x}) \mid \boldsymbol{x} \neq \boldsymbol{0}, \boldsymbol{x}=C\boldsymbol{y}, \boldsymbol{y} \in \hat{U}_k\}=\mu_k.\end{aligned} \tag{5.2.77}$$

(5.2.75)のもう一方の不等式も，最大・最小定理の(ii)の形を用いて証明できる．■

任意の $\boldsymbol{x} \in V_n (\boldsymbol{x} \neq \boldsymbol{0})$ に対して，$A\boldsymbol{x}$ になるべく近い $\mu\boldsymbol{x}$ は実数 $\mu=R_A(\boldsymbol{x})$ で与えられるという事実がある．これは，注意 3.7.5（いまの場合 $G=I_n$ としていることに注意)により，ベクトル $A\boldsymbol{x}$ を $\boldsymbol{x}$ の方向(の 1 次元部分空間)とそれに"直交"するベクトルとに分ける

$$A\boldsymbol{x}=\mu\boldsymbol{x}+\boldsymbol{z}, \qquad \boldsymbol{x}^{\mathrm{H}}\boldsymbol{z}=0 \tag{5.2.78}$$

という問題であるともいえる．しかるに(5.2.78)より直ちに，

$$\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}=\mu\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}, \quad \text{すなわち}, \quad \mu=R_A(\boldsymbol{x}) \tag{5.2.79}$$

が得られる．

注意 5.2.11　(5.2.78)において，$\boldsymbol{x}$ が A の固有ベクトルのどれかに一致したときには $\boldsymbol{z}=\boldsymbol{0}$ であるから，$\boldsymbol{x}$ が固有ベクトルから離れていればいるほど $\boldsymbol{z}$ は大きくなるはずである．しかし，

$$\sigma=\frac{\|\boldsymbol{z}\|_2}{\|\boldsymbol{x}\|_2} \tag{5.2.80}$$

とおくと，$\boldsymbol{x}(\in V_n)$をいろいろに変化させたとき，μ と σ は互いにまったく自由ではなく，μ を横軸，σ を縦軸に取ったとき A の固有値 $\lambda_1\geqq\cdots\geqq\lambda_n$ の最大 λ_1 と最小 λ_n を直径とする半円から，相続く固有値 λ_k,λ_{k+1} を直径とする $n-1$ 個の半円を除いた残りの領域(図 5.2.1)の部分にしか点(μ,σ)は存在しえないことが知られている(次の注意 5.2.12 を見よ)．これは，材料力学(弾性論)で古来有名な **Mohr**(モール)**の応力円**(Mohr's circle)の話の一般化であるとみなせる(A が**応力テンソル**(stress tensor)に対応)．$\boldsymbol{x}$ を法線とする断面の法線方向の応力が μ，剪断応力が σ に対応する．

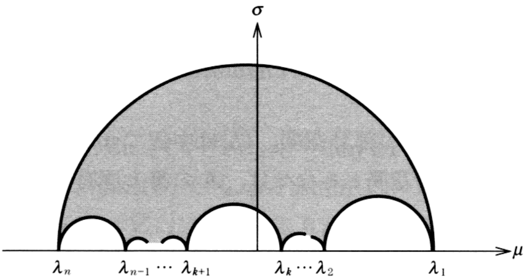


図 5.2.1　一般化された Mohr の応力円

注意 5.2.12　V_n の適当な基底を選んで，Rayleigh 商 $R_A(\boldsymbol{x})$を(5.2.62)のように表しておくと，上の注意 5.2.11 に述べたことは次のようにも言い直せる．簡単のため，一般性を失わずに $\|\boldsymbol{x}\|_2{}^2=\|\boldsymbol{y}\|_2{}^2=|y_1|^2+\cdots+|y_n|^2=1$ としておく．すると

$$\left.\begin{aligned}
\mu&=R_A(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}=\boldsymbol{x}^{\mathrm{H}}A^{\mathrm{H}}\boldsymbol{x}=\sum_{i=1}^{n}\lambda_i|y_i|^2,\\
\sigma^2&=\|\boldsymbol{z}\|_2{}^2=\|A\boldsymbol{x}-R_A(\boldsymbol{x})\,\boldsymbol{x}\|_2{}^2\\
&=\boldsymbol{x}^{\mathrm{H}}A^{\mathrm{H}}A\boldsymbol{x}-R_A(\boldsymbol{x})\,\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}-\boldsymbol{x}^{\mathrm{H}}A^{\mathrm{H}}\boldsymbol{x}R_A(\boldsymbol{x})+R_A(\boldsymbol{x})^2\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}\\
&=\boldsymbol{x}^{\mathrm{H}}A^2\boldsymbol{x}-\mu^2=\sum_{i=1}^{n}\lambda_i{}^2|y_i|^2-\mu^2.
\end{aligned}\right\} \tag{5.2.81}$$

いま，確率 $p_1=|y_1|^2,\cdots,p_n=|y_n|^2$ で値 $\lambda_1,\cdots,\lambda_n$ を取る確率変数 X(有限離散分布)を考えると，(5.2.81)は，μ が X の期待値(平均)$\mu=\mathrm{E}[X]$, σ が標準偏差 $\sigma=\sqrt{\mathrm{V}[X]}$ であることを表している．そこで，一般化された Mohr の円の話は，このような確率分布の平均と標準偏差の間の関係の話となる．平均が μ のとき，標準偏差あるいは分散を最大にする分布は，直観的には，なるべく離れた値(すなわち λ_1 と λ_n)を取る 2 点分布であると思われる．このような 2 点分布の平均 μ と標準偏差は $\sigma=\sqrt{(\lambda_1-\mu)(\mu-\lambda_n)}$ という，(μ,σ)平面上の半円で表される関係で結ばれる

ことはよく知られている．また，なるべく σ を小さくしようとすれば，μ を挟むなるべく近い 2 点 $\lambda_k, \lambda_{k+1}(\lambda_k \geqq \mu \geqq \lambda_{k+1})$ に値を取る 2 点分布を選べばよいであろう．このとき $\sigma=\sqrt{(\lambda_k-\mu)(\mu-\lambda_{k+1})}$ という関係がある．このように考えれば，注意 5.2.11 に述べた事実は“自然”であろう．より厳密な数学的な扱いについては，演習問題 5.9 を見よ．

注意 5.2.13 任意の非零ベクトル $\boldsymbol{x} \in V_n$ に対して $\mu=R_A(\boldsymbol{x})$ と $\sigma=\|A\boldsymbol{x}-\mu\boldsymbol{x}\|_2/\|\boldsymbol{x}\|_2$ が一般化された Mohr の応力円の指定する範囲にしか現れえないという事実は，Hermite 行列 A の固有値($\lambda_1 \geqq \cdots \geqq \lambda_n$ とする)の存在範囲について以下のような有用な情報を与えてくれる．

(i) 区間$[\mu-\sigma, \mu+\sigma]$に少なくとも 1 個の A の固有値 λ_i が存在する．[略証：$\lambda_1 \geqq \mu \geqq \lambda_n$ であるので，$\lambda_k \geqq \mu \geqq \lambda_{k+1}$ を満たす k が存在する．点(μ, σ)は$[\lambda_{k+1}, \lambda_k]$を直径とする半円の中にはないから $\sigma^2 \geqq (\lambda_k-\mu)(\mu-\lambda_{k+1})$．ゆえに，$0 \leqq \lambda_k-\mu \leqq \sigma$ あるいは $0 \leqq \mu-\lambda_{k+1} \leqq \sigma$(あるいはその両方)．]

(ii) μ を含む開区間$(\underline{\lambda}, \bar{\lambda})$が A の固有値を一つも含まないことが知られているとき，μ より大きい最小の A の固有値 λ_k と μ より小さい最大の A の固有値 λ_{k+1} とは

$$\mu-\frac{\sigma^2}{\bar{\lambda}-\mu} \leqq \lambda_{k+1} \leqq \mu, \qquad \mu \leqq \lambda_k \leqq \mu+\frac{\sigma^2}{\mu-\underline{\lambda}} \tag{5.2.82}$$

を満たす．[略証：$[\underline{\lambda}, \lambda_k]$を直径とする半円は$[\lambda_{k+1}, \lambda_k]$を直径とする半円に含まれるから，点$(\mu, \sigma)$はそれらの半円の外にあり $\sigma^2 \geqq (\lambda_k-\mu)(\mu-\underline{\lambda})$．ゆえに，$\mu \leqq \lambda_k \leqq \mu+\sigma^2/(\mu-\underline{\lambda})$．残りも同様．]

(iii) 開区間$(\underline{\lambda}, \bar{\lambda})$には μ と A の固有値 λ_i の 1 個だけ(重複するものはまとめて 1 個と数える)が含まれているということが知られているとき，

$$\mu-\frac{\sigma^2}{\bar{\lambda}-\mu} \leqq \lambda_i \leqq \mu+\frac{\sigma^2}{\mu-\underline{\lambda}}. \tag{5.2.83}$$

[略証：$\underline{\lambda}<\mu \leqq \lambda_i$ の場合と $\lambda_i \leqq \mu<\bar{\lambda}$ の場合に分けて考えれば(ii)とほぼ同様にして証明できる．それぞれ開区間$(\underline{\lambda}, \lambda_i), (\lambda_i, \bar{\lambda})$には A の固有値が存在しない．]

(iv) A の最大固有値 λ_1，最小固有値 λ_n の少なくとも一方は区間$(\mu-\sigma, \mu+\sigma)$の外にある．[略証：点(μ, σ)が$[\lambda_n, \lambda_1]$を直径とする円の中にあるから $\sigma^2 \leqq (\lambda_1-\mu)(\mu-\lambda_n)$．ゆえに，$\sigma \leqq \lambda_1-\mu$ (すなわち，$\mu+\sigma \leqq \lambda_1$)あるいは $\sigma \leqq \mu-\lambda_n$ (すなわち $\lambda_n \leqq \mu-\sigma$)，あるいはその両方．]

(v) 開区間$(\underline{\lambda}, \bar{\lambda})$が A の固有値をすべて含むということが知られているとき，

$$\underline{\lambda}<\lambda_n \leqq \mu-\frac{\sigma^2}{\bar{\lambda}-\mu}, \qquad \mu+\frac{\sigma^2}{\mu-\underline{\lambda}} \leqq \lambda_1<\bar{\lambda}. \tag{5.2.84}$$

[略証：(i)～(iv)と同様．]

5.2.5　固有値の摂動

行列 A の固有値や固有ベクトルが，A が変動を受けたときどの程度変動するかを見積ることは，応用上も非常に重要な問題である．ただ，この問題は，理論的にも数値計算的にも，一般にはかなり面倒な問題であって，多くの研究がなされてきたし，またなされつつある．最近やっとこの問題を中心に取り上げた成書が著されはじめた[1]．ここでは，この問題の一般論に深く立ち入ることはせず，ごく古典的な“やさしい”Hermite 行列の固有値の変動についてのみ簡単に触れる(→6.8.7 項でも関連するやや複雑な話題を取り上げる)．

A,B,C を n 次 Hermite 行列，それらの固有値を大きさの順に並べたものを $\lambda_1(A)\geqq\cdots\geqq\lambda_n(A)$，$\lambda_1(B)\geqq\cdots\geqq\lambda_n(B)$，$\lambda_1(C)\geqq\cdots\geqq\lambda_n(C)$ とする．

(i)　$A=B+C$ ならば，$i=1,\cdots,n$ に対して

$$\lambda_i(B)+\lambda_n(C)\leqq\lambda_i(A)\leqq\lambda_i(B)+\lambda_1(C), \tag{5.2.85}$$

$$|\lambda_i(A)-\lambda_i(B)|\leqq\|C\|_2=\|A-B\|_2. \tag{5.2.86}$$

特に C が非負定値なら

$$\lambda_i(B)\leqq\lambda_i(A). \tag{5.2.87}$$

[略証：5.2.4 項の最大・最小定理を用いる．$\boldsymbol{x}\neq\boldsymbol{0}$ に対して

$$\begin{aligned}&R_A(\boldsymbol{x})=R_B(\boldsymbol{x})+R_C(\boldsymbol{x}),\quad \lambda_n(C)\leqq R_C(\boldsymbol{x})\leqq\lambda_1(C),\\ &R_A(\boldsymbol{x})\leqq R_B(\boldsymbol{x})+\lambda_1(C),\quad R_A(\boldsymbol{x})\geqq R_B(\boldsymbol{x})+\lambda_n(C)\end{aligned} \tag{5.2.88}$$

である．U_i を i 次元部分空間として，$\lambda_i(A)=\max\limits_{U_i}\min\{R_A(\boldsymbol{x})\mid\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in U_i\}=\min\{R_A(\boldsymbol{x})\mid\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in\hat{U}_i\}$ と表すと，$\lambda_i(B)=\max\limits_{U_i}\min\{R_B(\boldsymbol{x})\mid\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in U_i\}\geqq\min\{R_B(\boldsymbol{x})\mid\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in\hat{U}_i\}$ であるから，(5.2.88) より (5.2.85) の右側の不等式を得る．また，$\lambda_i(A)=\min\limits_{U_{n-i+1}}\max\{R_A(\boldsymbol{x})\mid\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in U_{n-i+1}\}=\max\{R_A(\boldsymbol{x})\mid\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in\hat{U}_{n-i+1}\}$ と表せば，$\lambda_i(B)\leqq\max\{R_B(\boldsymbol{x})\mid\boldsymbol{x}\neq\boldsymbol{0},\boldsymbol{x}\in\hat{U}_{n-i+1}\}$ で，(5.2.85) の左側の不等式を得る．C が Hermite 行列であるから $\rho(C)=\|C\|_2=\max\{|\lambda_1(C)|,|\lambda_n(C)|\}$ であり，(5.2.86) が成り立つ．C が非負定値なら $\lambda_n(C)\geqq 0$ であるから (5.2.87) が成り立つ．]

(ii)　A,B が非負定値なら，$A+B$ も非負定値で(これは自明！)

$$\rho(A+B)\geqq\rho(A). \tag{5.2.89}$$

---

1)　たとえば，G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed., John Hopkins University Press, Baltimore, 1996；F. Chaitin-Chatelin, *Valeurs propres de matrices*, Masson, Paris, 1988［(邦訳)伊理正夫，伊理由美：「行列の固有値—最新の解法と応用」，シュプリンガー・フェアラーク東京，2003］；など．

[略証：スペクトル半径 ρ の定義と(5.2.87)より明らか.]

注意 5.2.14 一般の行列に対して摂動問題が厄介なのは次の例を見れば理解されよう.

$$A=\begin{bmatrix}2&1&0\\0&2&1\\0&0&2\end{bmatrix},\qquad B=\begin{bmatrix}2&1&0\\0&2&1\\\varepsilon&0&2\end{bmatrix} \tag{5.2.90}$$

において，A は 3 重固有値 $\lambda_1=2$ をもち，1 本の固有ベクトル $\boldsymbol{x}_1=[1,0,0]^\top$ と 2 本の一般化固有ベクトル $\boldsymbol{x}_2=[0,1,0]^\top, \boldsymbol{x}_3=[0,0,1]^\top$ をもち，それらは

$$A\boldsymbol{x}_3=\lambda_1\boldsymbol{x}_3+\boldsymbol{x}_2,\quad A\boldsymbol{x}_2=\lambda_1\boldsymbol{x}_2+\boldsymbol{x}_1,\quad A\boldsymbol{x}_1=\lambda_1\boldsymbol{x}_1 \tag{5.2.91}$$

を満たす．一方，B の特性多項式は

$$\varphi_B(x)=(2-x)^3+\varepsilon \tag{5.2.92}$$

なので，B は相異なる 3 個の固有値 $\mu_1=2+\varepsilon^{1/3}, \mu_2=2+\omega\varepsilon^{1/3}, \mu_3=2+\omega^2\varepsilon^{1/3}$ ($\varepsilon^{1/3}$ は ε の一つの 3 乗根，ω は 1 の原始 3 乗根)をもちそれぞれに対応する固有ベクトルは $\boldsymbol{x}_1=[1,\varepsilon^{1/3},\varepsilon^{2/3}]^\top, \boldsymbol{x}_2=[1,\omega\varepsilon^{1/3},\omega^2\varepsilon^{2/3}]^\top, \boldsymbol{x}_3=[1,\omega^2\varepsilon^{1/3},\omega\varepsilon^{2/3}]^\top$ である.

$$B-A=\begin{bmatrix}0&0&0\\0&0&0\\\varepsilon&0&0\end{bmatrix} \tag{5.2.93}$$

は，ε が小さいとき，どのようなノルムで計っても小さな行列なのに，固有値の変動は $\varepsilon^{1/3}$ とかなり大きくなり(もっともこれは重根を有する代数方程式に共通した現象であるが)，さらに，3 本の固有ベクトルは，線形独立であるとはいえ，かなり従属に近い．すなわち，B を対角化するための相似変換行列を $S=[\boldsymbol{x}_1,\boldsymbol{x}_2,\boldsymbol{x}_3]$ とすると，

$$\begin{bmatrix}2+\varepsilon^{1/3}&0&0\\0&2+\omega\varepsilon^{1/3}&0\\0&0&2+\omega^2\varepsilon^{1/3}\end{bmatrix}=S^{-1}BS,\quad S=\begin{bmatrix}1&1&1\\\varepsilon^{1/3}&\omega\varepsilon^{1/3}&\omega^2\varepsilon^{1/3}\\\varepsilon^{2/3}&\omega^2\varepsilon^{2/3}&\omega\varepsilon^{2/3}\end{bmatrix} \tag{5.2.94}$$

であり，$\det S=-3\sqrt{3}\,\mathrm{i}\varepsilon$ であるから S は特異に近い．また

$$S^{-1}=\frac{1}{3}\begin{bmatrix}1&\varepsilon^{-1/3}&\varepsilon^{-2/3}\\1&\omega^2\varepsilon^{-1/3}&\omega\varepsilon^{-2/3}\\1&\omega\varepsilon^{-1/3}&\omega^2\varepsilon^{-2/3}\end{bmatrix} \tag{5.2.95}$$

であるから，S の条件数(たとえば 2 乗ノルムに基づくもの(→ 注意 3.3.2))は

$$\mathrm{cond}_2\, S=\varepsilon^{-2/3} \tag{5.2.96}$$

となって，ε が小さいときには非常に大きくなる．(l 重固有値で類似の場合はどうなるか？ 注意 6.9.3 も見よ.)

5.3 Cayley-Hamilton の定理

注意 2.2.3 で述べたように，1 変数 m 次多項式 $f(x)=a_mx^m+\cdots+a_1x+a_0$ に対して n 次正方行列 A の行列多項式 $f(A)=a_mA^m+\cdots+a_1A+a_0I_n$(これも n 次正方行列)を定義することができる．A の固有値 λ に対応する固有ベクトルを $\boldsymbol{x}$ とすれば，$\boldsymbol{x}$ は $f(A)$ の固有値 $f(\lambda)$ に対応する固有ベクトルでもある．

特に $f(x)$ を A の特性多項式 $\varphi_A(x)$ に選ぶと

$$\varphi_A(A)=O_{n,n} \tag{5.3.1}$$

が成り立つ．(5.3.1)は **Cayley - Hamilton**(ケイリー - ハミルトン)**の定理**(Cayley と Hamilton の順は逆に並べられることもある)の名でよく知られている．これは次のようにして証明できる．

行列 $A-xI_n$(x は変数)の余因子行列(→2.7.3 項)を $G(x)=\mathrm{adj}(A-xI_n)$ と記すと，

$$(A-xI_n)G(x)=\det(A-xI_n)\cdot I_n=\varphi_A(x)I_n \tag{5.3.2}$$

が成り立つ(→(2.8.4))．ここで，

$$\varphi_A(x)=(-1)^nx^n+(-1)^{n-1}c_{n-1}x^{n-1}+\cdots-c_1x+c_0, \tag{5.3.3}$$

$$G(x)=(-1)^{n-1}C_{n-1}x^{n-1}+\cdots-C_1x+C_0 \tag{5.3.4}$$

と表しておく．(行列 $G(x)$ の各要素は $A-xI_n$ の $n-1$ 次の小行列式であるから，x のたかだか $n-1$ 次の多項式であり，その x^k の係数を集めて作った n 次正方行列が $(-1)^kC_k$ である．特に $C_0=\mathrm{adj}\,A$.)

(5.3.3)，(5.3.4)を(5.3.2)に代入して x^k の係数を比較すると，

$$\left.\begin{aligned}
AC_0 \qquad\qquad\qquad &= c_0I_n,\\
-AC_1-C_0 \qquad\qquad &= -c_1I_n,\\
\vdots \qquad\qquad\qquad &\\
(-1)^{n-1}AC_{n-1}-(-1)^{n-2}C_{n-2} &= (-1)^{n-1}c_{n-1}I_n,\\
-(-1)^{n-1}C_{n-1} &= (-1)^nI_n
\end{aligned}\right\} \tag{5.3.5}$$

を得る．(5.3.5)の各式の両辺に $I_n,A,A^2,\cdots,A^n$ を掛けて加えれば，左辺の和は $O_{n,n}$ となり，右辺の和は $\varphi_A(A)$ となる．

注意 5.3.1 上の定理から直ちに，$c_0=\det A\neq 0$ ならば $A^{-1}=(c_1I_n-c_2A+\cdots$

$+(-1)^{n-2}c_{n-1}A^{n-2}+(-1)^{n-1}A^{n-1})/c_0$ であることが示せる．

注意 5.3.2 一般に，任意の m 次多項式 $f(x)$ に対して

$$f(z)-f(y)=(y-z)g(y,z) \tag{5.3.6}$$

($g(y,z)$ は 2 変数 y,z の $m-1$ 次多項式)が成り立つから，y,z にそれぞれ A, xI_n を代入すれば，

$$f(xI)-f(A)=f(x)I-f(A)=(A-xI)G(x). \tag{5.3.7}$$

ここで，$G(x)=g(A,xI)$ は各要素が $m-1$ 次以下の x の多項式であるような n 次正方行列である．したがって，Cayley-Hamilton の定理の証明と同じ手続きによって，「ある多項式 $f(x)$ に対して $f(A)=O_{n,n}$ となるための必要十分条件は，

$$f(x)I=(A-xI)G(x) \tag{5.3.8}$$

を満たす(多項式を要素とする)行列 $G(x)$ が存在することである」ということもわかる．

注意 5.3.3 (5.3.5)をもっと詳しく分析してみる．最下式から k 個の式に下から順に $A^k, A^{k-1}, \cdots, A$ を掛けて加えると，

$$AC_{n-k}=(-1)^{k-1}(A^k-c_{n-1}A^{k-1}+c_{n-2}A^{k-2}-\cdots+(-1)^{k-1}c_{n-k+1}A)$$
$$(k=1,2,\cdots,n-1) \tag{5.3.9}$$

が得られる．$\varphi_A(x)=(-1)^n(x-\lambda_1)\cdots(x-\lambda_n)$ なら $\varphi_{A^p}(x)=(-1)^n(x-\lambda_1{}^p)\cdots(x-\lambda_n{}^p)$ であるから(→注意 5.2.2)，(5.1.17)より

$$\mathrm{tr}(AC_{n-k})=(-1)^{k-1}\left(\sum_{i=1}^n\lambda_i{}^k-c_{n-1}\sum_{i=1}^n\lambda_i{}^{k-1}+\cdots+(-1)^{k-1}c_{n-k+1}\sum_{i=1}^n\lambda_i\right). \tag{5.3.10}$$

ところで，基本対称式とべき対称式の間に成り立つ有名な **Newton**(ニュートン)**の公式**[1)]

$$kc_{n-k}=(-1)^{k-1}\left(\sum_{i=1}^n\lambda_i{}^k-c_{n-1}\sum_{i=1}^n\lambda_i{}^{k-1}+\cdots+(-1)^{k-1}c_{n-k+1}\sum_{i=1}^n\lambda_i\right)$$
$$(k=1,2,\cdots,n) \tag{5.3.11}$$

を想起すると，(5.3.10)は

$$\mathrm{tr}(AC_{n-k})=kc_{n-k} \tag{5.3.12}$$

と書き直せる．(5.3.5)と(5.3.12)を用いると，

---

1) この公式は多項式 $\psi(t)=\prod_{i=1}^n(1-t\lambda_i)=\sum_{k=0}^n(-1)^kt^kc_{n-k}$ ($c_n\equiv1$)を考え，$\psi'(t)=\sum_{k=1}^n(-1)^kkt^{k-1}c_{n-k}$，$\psi'(t)/\psi(t)=(\mathrm{d}/\mathrm{d}t)\log\psi(t)=-\sum_{i=1}^n\lambda_i/(1-t\lambda_i)=-\sum_{k=0}^\infty t^k\sum_{i=1}^n\lambda_i{}^{k+1}$ に注意して $\psi'(t)=\psi(t)(\mathrm{d}/\mathrm{d}t)\log\psi(t)$ の両辺の t^{k-1} ($k=1,\cdots,n$)の係数を比較すれば得られる．

$$
\left.\begin{aligned}
C_{n-1}&=I_n,\\
c_{n-1}&=\mathrm{tr}(AC_{n-1}),\\
C_{n-2}&=c_{n-1}I-AC_{n-1},\\
c_{n-2}&=\frac{1}{2}\mathrm{tr}(AC_{n-2}),\\
&\vdots\\
C_{n-k}&=c_{n-k+1}I_n-AC_{n-k+1},\\
c_{n-k}&=\frac{1}{k}\mathrm{tr}(AC_{n-k}),\\
&\vdots\\
C_1&=c_2I_n-AC_2,\\
c_1&=\frac{1}{n-1}\mathrm{tr}(AC_1),\\
C_0&=c_1I_n-AC_1,\\
c_0&=\frac{1}{n}\mathrm{tr}(AC_0),
\end{aligned}\right\} \tag{5.3.13}
$$

のように，逐次に固有値の基本対称式 c_k と $G(x)=\mathrm{adj}(A-xI)$ の x^k の係数行列 C_k とを計算して行くことができる．

このようにして，$\varphi_A(x)$ や $C_0=\mathrm{adj}\,A$ や $c_0=\det A$ が計算できるので，一時この計算法は **Frame**(フレーム)**法**とか **Faddeev**(ファデェーエフ)**法**とか **Leverrier-Faddeev**(ルヴリエ-ファデェーエフ)**法**とか呼ばれて，特性多項式や行列式 $c_0=\det A$ や逆行列 $A^{-1}=C_0/c_0$ を計算する"実際的な"方法として教えられていたこともあったが，(5.3.13)を実行するには行列積の計算を n 回おこなわなければならないので，総演算回数は n^4 に比例する．現在では，n^3 に比例する演算回数でそれらの量を計算する方法が種々開発されているので，この方法には歴史的意義と理論的興味が残るのみとなった．

注意 5.3.4 (5.3.13)をさらに続けると $C_{-1}=c_0I-AC_0$ は上記の性質により $C_{-1}=O$ となるので計算誤りの一つの検査法ともなる．

注意 5.3.5 Leverrier-Faddeev 法や Frame 法は，"除算を用いない($1/k$ は実質的な割り算ではない)"行列式(あるいは逆行列)の計算法として依然として理論的興味のあるものであり，特に，A が疎行列の場合には計算量的にも除算を含む各種の方法に対して競争力があるかもしれないとして，最近再び研究対象となっている．このような観点からの議論の一端を，本書でも補章 A.4 節，A.5 節で取り上げている．

5.4　最小消去多項式，最小多項式

本節の議論は，任意の体 K に対して成立する．また，すべての手続きは，K の上の有限回の有理演算によって実行できることにも注意されたい．

5.4.1　Krylov 列と巡回部分空間

あるベクトル $\boldsymbol{x}_0 \in V_n$ から出発して，線形変換 $A: V_n \to V_n$ を繰り返し施すことにより得られるベクトルの系列 $\boldsymbol{x}_0$，$A\boldsymbol{x}_0$，$A^2\boldsymbol{x}_0$，$\cdots$ (**Krylov**(クリロフ)**列** (Krylov sequence)と呼ばれる)を考える．$B=\{\boldsymbol{x}_0, A\boldsymbol{x}_0, \cdots, A^{p-1}\boldsymbol{x}_0\}$は線形独立であるが $A^p\boldsymbol{x}_0$ はB に線形従属であるというような p がある(V_n の中のベクトルであるから当然 $p \leqq n$)．このとき

$$A^p\boldsymbol{x}_0 = f_0\boldsymbol{x}_0 + f_1A\boldsymbol{x}_0 + \cdots + f_{p-1}A^{p-1}\boldsymbol{x}_0 \tag{5.4.1}$$

であるから，$A^{p+1}\boldsymbol{x}_0 = f_0A\boldsymbol{x}_0 + \cdots + f_{p-1}A^p\boldsymbol{x}_0 = f_0A\boldsymbol{x}_0 + \cdots + f_{p-2}A^{p-1}\boldsymbol{x}_0 + f_{p-1}(f_0\boldsymbol{x}_0 + \cdots + f_{p-1}A^{p-1}\boldsymbol{x}_0)$のように，$A^{p+1}\boldsymbol{x}_0$，$A^{p+2}\boldsymbol{x}_0$，$\cdots$はみな B に線形従属である．多項式

$$f(x) = x^p - f_{p-1}x^{p-1} - \cdots - f_1x - f_0 \tag{5.4.2}$$

を定義すれば，(5.4.1)は

$$f(A)\boldsymbol{x}_0 = \boldsymbol{0} \tag{5.4.3}$$

と書き表せる．また，B によって張られる部分空間を $L_A(\boldsymbol{x}_0) = L(\boldsymbol{x}_0, A\boldsymbol{x}_0, \cdots, A^{p-1}\boldsymbol{x}_0)$と記し，$\boldsymbol{x}_0$ からA によって生成される**巡回部分空間**(cyclic subspace)と呼ぶ．明らかに，巡回部分空間は A の不変部分空間(→3.3.4 項)である．その構成法から明らかなように，$L_A(\boldsymbol{x}_0)$は $\boldsymbol{x}_0$ を含む A の“最小の”不変部分空間である．

B は $L_A(\boldsymbol{x}_0)$の基底であるから，$A|_{L_A(\boldsymbol{x}_0)}$ は B に関して

$$C_A(\boldsymbol{x}_0) = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 & f_0 \\ 1 & 0 & \cdots & \cdots & 0 & f_1 \\ 0 & 1 & \ddots & & \vdots & f_2 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ \vdots & & \ddots & 1 & 0 & f_{p-2} \\ 0 & \cdots & \cdots & 0 & 1 & f_{p-1} \end{bmatrix} \tag{5.4.4}$$

という行列表現をもつ．この“形”の行列は**コンパニオン行列**(companion matrix)と呼ばれる(1 次のコンパニオン行列は$[f_0]$であるとする)．また，直接計算すればわかるように，

$$\varphi_{C_A(\boldsymbol{x}_0)}(x)=\det(C_A(\boldsymbol{x}_0)-xI_p)=(-1)^p f(x) \tag{5.4.5}$$

である．[$C_A(\boldsymbol{x}_0)-xI_p$ の行 p を x 倍して行 $p-1$ に加え，その行 $p-1$ を x 倍して行 $p-2$ に加え，……，行 2 を x 倍して行 1 に加えると，$(1,p)$要素に $-f(x)$ が作られる．そのとき，(p,p)以外の対角要素はみな 0 になっている．]

5.4.2　消去多項式

一般に，変数 x の最高次の係数が 1 であるような m 次多項式 $g(x)=x^m+\cdots$ (**モニック多項式**(monic polynomial)という)で $g(A)\boldsymbol{x}_0=\boldsymbol{0}$ を満たすものを，$\boldsymbol{x}_0$ $(\neq\boldsymbol{0})$の A に関する**消去多項式**(annihilator)と呼ぶ．$g(A)\boldsymbol{x}_0=\boldsymbol{0}$ は，$A^m\boldsymbol{x}_0$ が $\{A^{m-1}\boldsymbol{x}_0,\ \cdots,\ A\boldsymbol{x}_0,\ \boldsymbol{x}_0\}$に線形従属であることを示している．$\boldsymbol{x}_0$ から始まる A の Krylov 列に関して定義される(5.4.2)の $f(x)$ は $\boldsymbol{x}_0$ の消去多項式である．この $f(x)$の次数 p は，$\{\boldsymbol{x}_0,\ A\boldsymbol{x}_0,\ \cdots,\ A^{p-1}\boldsymbol{x}_0,\ A^p\boldsymbol{x}_0\}$がはじめて線形従属になるような，最小の p を与えているから，$\boldsymbol{x}_0$ の消去多項式の中で最小次数である．最小次数の消去多項式を**最小消去多項式**(minimal annihilator)と呼ぶ．コンパニオン行列(5.4.4)に関する $\boldsymbol{x}_0$ の最小消去多項式は $f(x)$である．

非零ベクトル $\boldsymbol{x}_0$ $(\in V_n)$に対してその最小消去多項式は一意に定まり，$\boldsymbol{x}_0$ のすべての消去多項式を割り切る．実際，m 次の $g(x)$が $\boldsymbol{x}_0$ の消去多項式であれば，$m\geqq p$ であるが，$g(x)$を $f(x)$で割ったときの商を $q(x)$，剰余を $r(x)$とすると，$r(x)$の次数は p より小さく，

$$g(x)=q(x)f(x)+r(x)\,. \tag{5.4.6}$$

$g(A)\boldsymbol{x}_0=\boldsymbol{0}$，$f(A)\boldsymbol{x}_0=\boldsymbol{0}$ より $r(A)\boldsymbol{x}_0=\boldsymbol{0}$ となるから，$r(x)=0$ でなければ $r(x)$(をモニック化したもの)が $\boldsymbol{x}_0$ の消去多項式となって $f(x)$の最小次数性に反する．($m=p$ のときには $q(x)=1$ となるので $g(x)=f(x)$．このことから最小消去多項式が一意に定まることがわかる．)

注意 5.4.1　$\boldsymbol{x}_0$ の最小消去多項式 $f(x)$が互いに素な(モニック)多項式 $f^{(1)}(x)$，$f^{(2)}(x)$の積に表されるとき，すなわち

$$f(x)=f^{(1)}(x)\cdot f^{(2)}(x) \tag{5.4.7}$$

であるとき，$f^{(i)}(x)$を最小消去多項式にもつベクトル $\boldsymbol{x}_i(\neq\boldsymbol{0})\in L_A(\boldsymbol{x}_0)$が存在して，直和分解

$$L_A(\boldsymbol{x}_0)=L_A(\boldsymbol{x}_1)\oplus L_A(\boldsymbol{x}_2) \tag{5.4.8}$$

が成り立つ．

証明　$f^{(1)},f^{(2)}$は互いに素であるから，それらの最大公約多項式は1に等しく，よく知られた代数学の定理により，

$$f^{(1)}(x)g^{(1)}(x)+f^{(2)}(x)g^{(2)}(x)=1 \tag{5.4.9}$$

を満たす多項式 $g^{(1)}(x),g^{(2)}(x)$が存在する($i\neq j$ なら $f^{(i)}(x)$と $g^{(j)}(x)$は互いに素)．(5.4.9)は

$$f^{(1)}(A)g^{(1)}(A)+f^{(2)}(A)g^{(2)}(A)=I \tag{5.4.10}$$

も意味する．

$$P_1=f^{(2)}(A)g^{(2)}(A),\qquad P_2=f^{(1)}(A)g^{(1)}(A) \tag{5.4.11}$$

とおくと

$$P_1+P_2=I. \tag{5.4.12}$$

任意の $\boldsymbol{x}\in L_A(\boldsymbol{x}_0)$に対して，$f(A)\boldsymbol{x}=\boldsymbol{0}$ であるから

$$P_1P_2\boldsymbol{x}=P_2P_1\boldsymbol{x}=\boldsymbol{0}. \tag{5.4.13}$$

(5.4.12)と(5.4.13)とより

$$P_i^2\boldsymbol{x}=P_i\boldsymbol{x} \tag{5.4.14}$$

が得られる．このことは，P_i が $L_A(\boldsymbol{x}_0)$における射影であって $L_A(\boldsymbol{x}_0)=L_1\oplus L_2$ $(L_i=P_i\cdot L_A(\boldsymbol{x}_0))$と直和分解されることを意味している(→注意3.7.3)．特に，

$$\boldsymbol{x}_1=P_1\boldsymbol{x}_0,\qquad \boldsymbol{x}_2=P_2\boldsymbol{x}_0 \tag{5.4.15}$$

とおけば，

$$f^{(1)}(A)\boldsymbol{x}_1=f^{(1)}(A)P_1\boldsymbol{x}_0=f^{(1)}(A)f^{(2)}(A)g^{(2)}(A)\boldsymbol{x}_0=g^{(2)}(A)f(A)\boldsymbol{x}_0=\boldsymbol{0} \tag{5.4.16}$$

であるから，$f^{(1)}(x)$は $\boldsymbol{x}_1$ の消去多項式である．さらに，$\boldsymbol{x}_1$ の任意の消去多項式を $h^{(1)}(x)$とすれば，

$$h^{(1)}(A)\boldsymbol{x}_1=h^{(1)}(A)P_1\boldsymbol{x}_0=h^{(1)}(A)f^{(2)}(A)g^{(2)}(A)\boldsymbol{x}_0=\boldsymbol{0} \tag{5.4.17}$$

であるから，$h^{(1)}(x)f^{(2)}(x)g^{(2)}(x)$が $f(x)=f^{(1)}(x)f^{(2)}(x)$で割り切れなければならない．$f^{(1)}$と $f^{(2)},g^{(2)}$が素だから，$h^{(1)}$は $f^{(1)}$で割り切れる．すなわち $f^{(1)}(x)$は $\boldsymbol{x}_1$ の最小消去多項式である．$L_1=L_A(\boldsymbol{x}_1)$，$L_2=L_A(\boldsymbol{x}_2)$についても同様．■

$f(x)$が3個以上の互いに素な因子の積に表される場合についても，順々に上の議論を繰り返すことによって，同様の結論が得られる．

5.4.3　最小多項式

V_n のすべてのベクトル $\boldsymbol{x}(\neq\boldsymbol{0})$の A に関する消去多項式であって，かつ次数

が最小のものを $\psi_A(x)$ と記して，A の**最小多項式**(minimal polynomial)と呼ぶ．$\psi_A(x)$ はどのベクトル $\boldsymbol{x}_0$ の最小消去多項式でも割り切られる．すべての $\boldsymbol{x} \in V_n$ を消去するためには $\psi_A(A)=O$ でなければならない．5.4.2 項における最小消去多項式に関する議論と同様にして，

- $\psi_A(x)$ は A の特性多項式 $\varphi_A(x)$ (これは $(-1)^n$ 倍するとモニックになる)を割り切る

ことがわかる．注意 5.4.1 と同様の議論をすれば，

- $\psi_A(x)=\psi_A^{(1)}(x)\cdot\cdots\cdot\psi_A^{(s)}(x)$ ($\psi_A^{(1)}(x),\cdots,\psi_A^{(s)}(x)$ は互いに素)という因子分解に対応して，$V_n=L_1\oplus\cdots\oplus L_s$ (L_i は V_n の A に関する不変部分空間)という直和分解がある

こともわかる．

注意 5.4.2 V_n の特定のベクトル $\boldsymbol{x}_0$ の最小消去多項式 $f(x)$ が $\psi_A(x)$ を割り切ることはもちろんである．$\psi_A(x)$ は有限個の因子しかもたず(個数は体 K の選び方によって異なるが)，次数が $\psi_A(x)$ の次数より小さい各因子 $g(x)$ に対しては $g(A) \neq O$ であるから，$g(A)\boldsymbol{x}=\boldsymbol{0}$ を満たす $g(x)$ を最小消去多項式としてもつ $\boldsymbol{x}$ の全体は(存在したとしても) V_n の真部分空間をなす．そこで，A の最小多項式 $\psi_A(x)$ は，そのような有限個の真部分空間に属するもの以外の“ほとんどすべての”ベクトル $\boldsymbol{x}$ の最小多項式となっている．

もっと明示的に述べると，K を超越拡大して，“一般の”ベクトル $\boldsymbol{x} \in V_n$ に対して $\boldsymbol{x}=[1,\varepsilon,\varepsilon^2,\cdots,\varepsilon^{n-1}]^\top$ (ε は超越元，$K=\mathbb{R}$ あるいは $\mathbb{C}$ のときには“無限小”とみなせばよい)とおき，それから出発する Krylov 列を考えると，

$$A^p\boldsymbol{x}=c_{p-1}A^{p-1}\boldsymbol{x}+\cdots+c_1A\boldsymbol{x}+c_0\boldsymbol{x} \tag{5.4.18}$$

は結局

$$A^p=c_{p-1}A^{p-1}+\cdots+c_1A+c_0I \tag{5.4.19}$$

を意味する．このことは，A の最小多項式 $\psi_A(x)$ が「$\psi_A(A)=O$ を満たす最小次数のモニック多項式」としても定義できるということを示している．

注意 5.4.3 A の最小多項式 $\psi_A(x)$ は特性多項式 $\varphi_A(x)$ を割り切るが，$\varphi_A(x)$ の零点(A の固有値)はすべて $\psi_A(x)$ の零点に含まれる．(実際，A の固有値 λ に対応する固有ベクトルを $\boldsymbol{x}(\neq\boldsymbol{0})$ とすれば，$\psi_A(x)$ は $\boldsymbol{x}$ の消去多項式でもあるから $\psi_A(A)\boldsymbol{x}=\boldsymbol{0}$．$A\boldsymbol{x}=\lambda\boldsymbol{x}$ であるから $\psi_A(\lambda)\boldsymbol{x}=\boldsymbol{0}$．$\boldsymbol{x}\neq\boldsymbol{0}$ であるから $\psi_A(\lambda)=0$.)すなわち，多重度を除いて $\varphi_A(x)$ の零点と $\psi_A(x)$ の零点は一致する．したがって，A の固有値の代数的重複度がすべて 1 ($\varphi_A(x)$ の零点がすべて単純) ならば $\varphi_A(x)=(-1)^n\psi_A(x)$．

5.4.4 $A-xI$の単因子

2.15 節(特に注意 2.15.24，2.15.25)における多項式環 $K[x]$の上での行列の行列式因子，単因子の話を，n 次正方行列 A から作られる行列 $A-xI$ に適用してみよう．すると，$\det S(x) \in K-\{0\}$，$\det T(x) \in K-\{0\}$を満たす x の多項式を要素とする行列(すなわち $K[x]$の可逆元)$S(x)$，$T(x)$が存在して

$$S(x)(A-xI)T(x)=\begin{bmatrix} e_1(x) & & & O \\ & e_2(x) & & \\ & & \ddots & \\ O & & & e_n(x) \end{bmatrix} \tag{5.4.20}$$

となる($\det S(x)$，$\det T(x)$は 1 に等しいとして一般性を失わない)．ここで，多項式 $e_i(x)$(**単因子**)は $e_{i+1}(x)$を割り切る$(i=1,\cdots,n-1)$．また，$A-xI$ の k 次の小行列式全体の最大公約多項式(**行列式因子**)(モニック多項式とする)を $d_k(x)$ と記すと$(d_0(x)\equiv 1)$，$e_k(x)=d_k(x)/d_{k-1}(x)$，$\varphi_A(x)=\det(A-xI)=(-1)^n d_n(x)=(-1)^n e_1(x)\cdot e_2(x)\cdot\cdots\cdot e_n(x)$である．

$A-xI$ の単因子について調べるために次の技術的な準備をする．

補助定理 与えられた n 次正方行列 A，および $K[x]$の上の n 次正方行列 $P(x)$に対して，適当な $K[x]$の上の行列 $C(x)$，$D(x)$と定数行列 S，T を選んで

$$P(x)=(A-xI)C(x)+S \quad \text{および} \quad P(x)=D(x)(A-xI)+T \tag{5.4.21}$$

と表すことができる．

証明 $P(x)=x^m P_m+x^{m-1}P_{m-1}+x^{m-2}P_{m-2}+\cdots+xP_1+P_0$，$C(x)=x^{m-1}C_{m-1}+x^{m-2}C_{m-2}+\cdots+xC_1+C_0$ (P_k, C_k は n 次正方定数行列)とおけば，(5.4.21)の第 1 式の両辺の x^k $(k=m,\ m-1,\ \cdots,\ 1,\ 0)$の係数を比較することによって，$P_m=-C_{m-1}$，$P_{m-1}=AC_{m-1}-C_{m-2}$，$\cdots$，$P_k=AC_k-C_{k-1}$，$\cdots$，$P_1=AC_1-C_0$，$P_0=AC_0+S$ により，$P(x)$と A から C_{m-1}，C_{m-2}，$\cdots$，C_1，C_0，S をこの順に次々と定めることができる($P(x)$，A から $C(x)$，S は一意に定まる)．第 2 式についても同様． ∎

$A-xI$ は $K[x]$の上の単模変換(行列式が定数(一般性を失わずに 1 であるとしてよい)の多項式行列，すなわち可逆行列を右と左から掛ける変換)で(5.4.20)

の形にできるが，他の行列 B について $B-xI$ も同じ形にできる．もし両者の単因子が一致するならば，$A-xI$ と $B-xI$ が単模変換で結ばれることになる(これは単模変換の可逆性から明らか)．ところが，$A-xI$ という形の行列に限ると，$A-xI$ と $B-xI$ の単模同値性と A と B の相似性とが等価になるという興味深い事実がある．

定理 5.4.1　n 次正方行列 A,B に対して，$K[x]$の上の(可逆)行列 $P(x)$，$Q(x)$が存在して

$$A-xI=P(x)(B-xI)Q(x) \tag{5.4.22}$$

であるための必要十分条件は，A と B が相似である，すなわち，正則な n 次正方行列 S が存在して

$$B=S^{-1}AS \tag{5.4.23}$$

となることである．

証明　十分性は明らか($(A-xI)=S(B-xI)S^{-1}$ で，S,S^{-1} は自明に $K[x]$ の上の可逆行列)．

必要性を示す．上の補助定理により

$$\begin{aligned} P(x)&=(A-xI)C(x)+S,\\ P(x)^{-1}&=(B-xI)D(x)+F,\\ Q(x)^{-1}&=G(x)(B-xI)+T \end{aligned} \tag{5.4.24}$$

と表しておく．(5.4.22)を

$$(A-xI)Q(x)^{-1}=P(x)(B-xI) \tag{5.4.25}$$

と書き換えておいてから，(5.4.24)を代入すると，

$$(A-xI)[G(x)(B-xI)+T]=[(A-xI)C(x)+S](B-xI).$$

括弧をほどいて移項すると

$$(A-xI)[G(x)-C(x)](B-xI)=S(B-xI)-(A-xI)T \tag{5.4.26}$$

を得る．この右辺は x について 1 次なのに，左辺は$[\cdots]$が零行列でない限り x について 2 次以上である．そこで，左辺の$[\cdots]$は零行列でなければならない．したがって右辺も零行列で，

$$S(B-xI)=(A-xI)T \quad \text{すなわち} \quad SB=AT,\ S=T \tag{5.4.27}$$

が得られる．$S=T$ が正則なら $SB=AT$ は(5.4.23)に他ならない．

S の正則性は，(5.4.24)の $P(x)$，$P(x)^{-1}$ の表式と $I=P(x)P(x)^{-1}$ より次

のようにして示せる．すなわち

$$I=P(x)P(x)^{-1}=[(A-xI)C(x)+S][(B-xI)D(x)+F]$$
$$=(A-xI)C(x)[(B-xI)D(x)+F]+S(B-xI)D(x)+SF$$

に(5.4.27)の第1式を代入すると

$$I=(A-xI)\{C(x)[(B-xI)D(x)+F]+T\cdot D(x)\}+SF \tag{5.4.28}$$

となるが，(5.4.28)の右辺は{…}が零行列でなければ x について1次以上である．しかるに左辺は定数，すなわち x について0次であるから，{…}は零行列でなければならない．このことから，$I=SF$，すなわち S が正則でなければならない． ∎

例 5.4.1　n 次コンパニオン行列(→(5.4.4))

$$C=\begin{bmatrix} 0 & \cdots & \cdots & 0 & f_0 \\ 1 & 0 & & \vdots & f_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & f_{n-2} \\ 0 & \cdots & 0 & 1 & f_{n-1} \end{bmatrix} \tag{5.4.29}$$

に対して，$C-xI$ の行列式因子は，

$$\left.\begin{aligned} d_n(x)&=(-1)^n\varphi_C(x)=x^n-f_{n-1}x^{n-1}-\cdots-f_1x-f_0, \\ d_k(x)&=1 \qquad (k=1,2,\cdots,n-1) \end{aligned}\right\} \tag{5.4.30}$$

である($k\leqq n-1$ に対しては，$C-xI$ の左下に k 次の単位行列があるので $d_k(x)=1$)．したがって，$C-xI$ の単因子は

$$e_1(x)=e_2(x)=\cdots=e_{n-1}(x)=1, \quad e_n(x)=(-1)^n\varphi_C(x) \tag{5.4.31}$$

であり，C の最小多項式は $e_n(x)$ である． □

例 5.4.2　(5.2.31c)の形の n 次正方行列

$$J=\begin{bmatrix} \lambda & 1 & & & \\ & \lambda & \ddots & O & \\ & & \ddots & \ddots & \\ O & & & \ddots & 1 \\ & & & & \lambda \end{bmatrix} \tag{5.4.32}$$

に対して，$J-xI$ の行列式因子は，

$$\left.\begin{aligned} d_n(x)&=(-1)^n\varphi_J(x)=(x-\lambda)^n, \\ d_k(x)&=1 \qquad (k=1,\cdots,n-1) \end{aligned}\right\} \tag{5.4.33}$$

であり($d_k(x)$ $(k\leqq n-1)$については $J-xI$ の右上の部分に着目せよ)，単因子は

$$e_1(x)=\cdots=e_{n-1}(x)=1, \qquad e_n(x)=(x-\lambda)^n \tag{5.4.34}$$

である． □

例 5.4.3　A の固有値がすべて異なるときには $\varphi_A(x)=(-1)^n\prod_{i=1}^{n}(x-\lambda_i)$ ($\lambda_1,\cdots,\lambda_n$ はすべて異なる)と $e_i(x)|e_{i+1}(x)$ と $(-1)^n\varphi_A(x)=\prod_{i=1}^{n}e_i(x)$ とにより，$e_1(x)=\cdots=e_{n-1}(x)=1, e_n(x)=(-1)^n\varphi_A(x)$ でなければならないことがわかる．

A の特性多項式 $\varphi_A(x)$ が K の上で既約であるときにも同様である． □

注意 5.4.4　n 次正方行列 A に対して $A-xI$ の単因子が(5.4.20)で与えられているとする．$e_i(x)|e_{i+1}(x)$ で $e_1(x)e_2(x)\cdots e_n(x)=(-1)^n\varphi_A(x)$ が x について n 次であるから，普通はかなり多くの $e_1(x),\cdots,e_k(x)$ が 1 に等しい．1 に等しくない単因子を次数の大きさの昇順に並べて，あらためて $e_1(x),\cdots,e_t(x)$ と記す．それぞれ x について $p_1,\cdots,p_t$ 次であるとする($p_1+\cdots+p_t=n$；$p_1\leqq\cdots\leqq p_t$)．各 $(-1)^{p_i}e_i(x)$ を特性多項式としてもつコンパニオン行列(p_i 次正方)を

$$C_i=\begin{bmatrix} 0 & \cdots & \cdots & 0 & f_0^{(i)} \\ 1 & \ddots & & \vdots & \vdots \\ & \ddots & \ddots & \vdots & \vdots \\ & & \ddots & 0 & \vdots \\ & O & & 1 & f_{p_i-1}^{(i)} \end{bmatrix} \tag{5.4.35}$$

とし，n 次正方行列

$$C=C_1\dot{+}\cdots\dot{+}C_t \tag{5.4.36}$$

を定義する．各 C_i は適当な $K[x]$ の上の可逆行列 $P_i(x)$，$Q_i(x)$ により

$$P_i(x)(C_i-xI)Q_i(x)=\begin{bmatrix} 1 & & & & \\ & \ddots & & O & \\ & & 1 & & \\ & O & & & \\ & & & & e_i(x) \end{bmatrix}\equiv E_i(x) \tag{5.4.37}$$

と変換できるので，$P(x)=P_1(x)\dot{+}\cdots\dot{+}P_t(x)$，$Q(x)=Q_1(x)\dot{+}\cdots\dot{+}Q_t(x)$ とおくと

$$P(x)(C-xI)Q(x)=E_1(x)\dot{+}\cdots\dot{+}E_t(x). \tag{5.4.38}$$

(5.4.38)の右辺は，対角行列で，対角要素の順を並べ換えれば(5.4.20)に一致するので，上記の定理 5.4.1 により，A は C に相似であることがわかる．

このように，任意の正方行列 A は相似変換によって(5.4.36)，(5.4.35)の形に変換することができる．ここで，$\varphi_{C_i}(x)$ は $A-xI$ の 1 と異なる単因子に(符号を除いて)等しい．C の最大ブロック C_t の最小多項式 $\psi_{C_t}(x)$ は特性多項式 $\varphi_{C_t}(x)$ に等しく，$A-xI$ の第 n 単因子に等しい．$\varphi_{C_t}(C_t)=O_{p_t,p_t}$ であるが，$\varphi_{C_i}(x)|\varphi_{C_t}(x)$ であるから，$\varphi_{C_t}(C_i)=O_{p_i,p_i}$．したがって，$\varphi_{C_t}(C)=O_{n,n}$．このことは，$C$ の最小多項式が $\varphi_{C_t}(x)$ であることを意味している．このようにして，$A-xI$ の第 n 単因子が A の最小多項式 $\psi_A(x)$ に等しいことがわかる．

注意 5.4.5 $A-xI$ の単因子を計算して(5.4.20)の形を定める，あるいは(5.4.38)の形を定めるには，手間はかなりかかりはするが，K 上の有理演算だけで十分である(固有値そのものの計算は必要ない)ことに注意すべきである．

5.5 Kronecker 積の固有値

m 次正方行列 A と n 次正方行列 B の Kronecker 積 $A\otimes B$(→2.4.2 項)は mn 次正方行列であるが，A の固有値を $\lambda_1,\cdots,\lambda_m$，$B$ の固有値を $\mu_1,\cdots,\mu_n$ とすると $A\otimes B$ の固有値は $\lambda_i\mu_j$ $(i=1,\cdots,m\,;\,j=1,\cdots,n)$ である．実際，A の Schur 形

$$R_1=Q_1{}^{\mathrm{H}}AQ_1 \tag{5.5.1}$$

(Q_1 は m 次ユニタリ；R_1 の対角要素は $\lambda_1,\cdots,\lambda_m$)と B の Schur 形

$$R_2=Q_2{}^{\mathrm{H}}BQ_2 \tag{5.5.2}$$

(Q_2 は n 次ユニタリ；R_2 の対角要素は $\mu_1,\cdots,\mu_n$)を用いると，(2.4.6)により

$$(Q_1{}^{\mathrm{H}}\otimes Q_2{}^{\mathrm{H}})(A\otimes B)(Q_1\otimes Q_2)=R_1\otimes R_2 \tag{5.5.3}$$

であるが，(2.4.5)により三角行列 R_1，R_2 の Kronecker 積は三角行列であるし，$(Q_1{}^{\mathrm{H}}\otimes Q_2{}^{\mathrm{H}})(Q_1\otimes Q_2)=I_m\otimes I_n=I_{mn}$ であるから $Q_1\otimes Q_2$ は mn 次ユニタリであり，$R_1\otimes R_2$ は $A\otimes B$ の Schur 形である．その mn 個の対角要素が $\lambda_i\mu_j$ であることは明らか．

同様な考え方で，$K=\mathbb{C}$ で A も B も Hermite 行列，それぞれの符号指数が $m_+,m_-,m_0\,;\,n_+,n_-,n_0$ であれば，$A\otimes B$ の符号指数は $(m_+\cdot n_+)+(m_-\cdot n_-)$，$(m_+\cdot n_-)+(m_-\cdot n_+)$，$(m_++m_-)\cdot n_0+m_0\cdot(n_++n_-)+m_0\cdot n_0$ である．

さらに，K 上の 2 変数の多項式 $\varphi(x,y)=\sum_{i,j}c_{ij}x^iy^j$ $(c_{ij}\in K)$ に対して

$\varphi(A;B)=\sum_{ij}c_{ij}(A^i\otimes B^j)$と定義すると，$\varphi(A;B)$の固有値は$m\cdot n$個の値$\varphi(\lambda_\alpha,\mu_\beta)$ $(\alpha=1,\cdots,m;\beta=1,\cdots,n)$である．ここで，$\lambda_1,\cdots,\lambda_m$は$A$の固有値，$\mu_1,\cdots,\mu_n$は$B$の固有値である．特に$\varphi(x,y)=x-y$とおくと$\varphi(A;B)=A\otimes I_n-I_m\otimes B$となるが(→演習問題4.4)，$A$と$B$の特性多項式を，それぞれ，

$$\left.\begin{aligned} f(x)&=(-x)^m+a_{m-1}(-x)^{m-1}+\cdots-a_1x+a_0,\\ g(x)&=(-x)^n+b_{m-1}(-1)^{n-1}+\cdots-b_1x+b_0 \end{aligned}\right\} \tag{5.5.4}$$

とすると，$\varphi(A;B)$の固有値は$\lambda_\alpha-\mu_\beta$ $(\alpha=1,\cdots,m;\beta=1,\cdots,n)$である．そこで$\det\varphi(A;B)=\det(A\otimes I_n-I_m\otimes B)=\prod_{\substack{\alpha=1,\cdots,m\\ \beta=1,\cdots,n}}(\lambda_\alpha-\mu_\beta)$は$f(x)$と$g(x)$の終結式(→例4.3.1)となることがわかる．

5.6　値　　　域

$K=\mathbb{C}$とする．Hermite行列に関するRayleigh商の一般化として，任意のn次正方行列Aに対して

$$\frac{\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}} \tag{5.6.1}$$

という量を考えることができる．$\boldsymbol{x}$を変えたとき($\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=1$として一般性を失わない)，この量が複素平面上で取る値の集合(有界閉集合；後に示すように凸集合でもある)をAの**値域**(または**値集合**；value field, field of values[1])と呼び$F(A)$と記す．

$$F(A)=\{\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}\mid \boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=1\} \tag{5.6.2}$$

$\lambda,\boldsymbol{x}$をAの一つの固有値とそれに対応する固有ベクトルとすると，$A\boldsymbol{x}=\lambda\boldsymbol{x}$であるから$\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}=\lambda\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}$であり，

$$\lambda_i\in F(A)\qquad(\lambda_1,\cdots,\lambda_n：A\text{の固有値}), \tag{5.6.3}$$

すなわち，値域はすべての固有値を含む．

特に，Aが正規行列であれば，ユニタリ相似変換で対角化されるから(→5.2.2項)，$Q^{\mathrm{H}}AQ=\mathrm{diag}[\lambda_1,\cdots,\lambda_n]$，$\boldsymbol{x}=Q\boldsymbol{y}$ $(\boldsymbol{y}^{\mathrm{H}}\boldsymbol{y}=1)$とおけば

$$\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}=\sum_{i=1}^n|y_i|^2\lambda_i,\quad |y_i|^2\geqq 0,\quad \sum_{i=1}^n|y_i|^2=1 \tag{5.6.4}$$

1)　関数の定義域(domain)と値域(range)というときの値域の一種ではあるが，かなり限定された意味で用いられている．(英語では呼び名は区別されている．)

である．(5.6.4)は，$F(A)$がAの固有値$\lambda_1,\cdots,\lambda_n$の**凸包**(convex hull)($\lambda_1,\cdots,\lambda_n$を含む最小の閉凸集合)であることを表している．

例 5.6.1　n次置換行列

$$P_n=\begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & \ddots & 1 \\ 1 & 0 & \cdots & \cdots & 0 \end{bmatrix}$$

は直交行列で($P^{\mathrm{T}}P=I_n$)，特性方程式$\varphi_{P_n}(x)=(-x)^n-(-1)^n=(-1)^n(x^n-1)$ $=(-1)^n\prod_{i=0}^{n-1}\left(x-\exp\left(\mathrm{i}\dfrac{2\pi i}{n}\right)\right)$をもつ．固有値$\lambda_1,\cdots,\lambda_n$は1の$n$乗根である．そこで，$P_n$の値域$F(P_n)$は，複素平面上の単位円に内接し，一つの頂点を1にもつ正n角形である．　□

例 5.6.2　Hermite 行列Aの値域$F(A)$は，その最大固有値λ_1と最小固有値λ_nを結ぶ実数軸上の直線分である．　□

$F(A)$が，一般に，凸集合であることを示すには，次のようにすればよい．まず，任意の複素数$\alpha(\in\mathbb{C})$に対して

$$F(\alpha A)=\{\alpha z\mid z\in F(A)\}=\alpha F(A) \tag{5.6.5}$$

であることに注意する．$F(A)$が凸であるということは，任意の二つの複素数$z^{(1)}$，$z^{(2)}\in F(A)$に対して，それらを結ぶ直線分$\{(1-t)z^{(1)}+tz^{(2)}\mid 0\leqq t\leqq 1\}$が$F(A)$に含まれることである．$F(A)$が凸であることと$F(\alpha A)$が凸であることは同じことであるから，(5.6.5)に注意すれば$\mathrm{Re}\,z^{(1)}=\mathrm{Re}\,z^{(2)}$であると仮定してよい．

行列Aは，二つの Hermite 行列A_{R}，A_{I}を用いて，

$$A=A_{\mathrm{R}}+\mathrm{i}A_{\mathrm{I}},\quad A_{\mathrm{R}}{}^{\mathrm{H}}=A_{\mathrm{R}},\quad A_{\mathrm{I}}{}^{\mathrm{H}}=A_{\mathrm{I}} \tag{5.6.6}$$

と一意に表せる．V_nの正規直交基底を適当に選ぶことによって，すなわち適当なユニタリ変換によって，

$$\left.\begin{aligned} \boldsymbol{x}^{\mathrm{H}}A_{\mathrm{R}}\boldsymbol{x}&=\lambda_1|x_1|^2+\cdots+\lambda_n|x_n|^2, \\ \boldsymbol{x}^{\mathrm{H}}\cdot\boldsymbol{x}&=|x_1|^2+\cdots+|x_n|^2=1 \end{aligned}\right\} \tag{5.6.7}$$

と表せる($\lambda_1,\cdots,\lambda_n$は$A_{\mathrm{R}}$の固有値)．値$z^{(1)}$，$z^{(2)}$を与えるベクトルを$\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(2)}$

とする：

$$z^{(k)}=\boldsymbol{x}^{(k)\mathrm{H}}A\boldsymbol{x}^{(k)}, \qquad \boldsymbol{x}^{(k)\mathrm{H}}\cdot\boldsymbol{x}^{(k)}=1 \qquad (k=1,2). \tag{5.6.8}$$

$\mathrm{Re}\, z^{(1)}=\mathrm{Re}\, z^{(2)}$ と(5.6.6)により

$$\left.\begin{aligned}\lambda_1|x_1{}^{(1)}|^2+\cdots+\lambda_n|x_n{}^{(1)}|^2&=\lambda_1|x_1{}^{(2)}|^2+\cdots+\lambda_n|x_n{}^{(2)}|^2,\\ |x_1{}^{(1)}|^2+\cdots+|x_n{}^{(1)}|^2&=|x_1{}^{(2)}|^2+\cdots+|x_n{}^{(2)}|^2=1.\end{aligned}\right\} \tag{5.6.9}$$

少々天下り的であるが，

$$x_i{}^{(k)}=r_i{}^{(k)}\mathrm{e}^{\mathrm{i}\theta_i{}^{(k)}} \qquad (r_i{}^{(k)}(\geqq 0),\ \theta_i{}^{(k)}:\text{実数}\,;\,k=1,2\,;\,i=1,\cdots,n) \tag{5.6.10}$$

とおいて，$\boldsymbol{x}(t)$を

$$\left.\begin{aligned}&x_i(t)=r_i(t)\mathrm{e}^{\mathrm{i}\theta_i(t)},\\ &r_i(t)=\sqrt{(1-t)\,r_i{}^{(1)2}+tr_i{}^{(2)2}}, \qquad \theta_i(t)=(1-t)\,\theta_i{}^{(1)}+t\theta_i{}^{(2)}\end{aligned}\right\} \tag{5.6.11}$$

により定義する．明らかに $\boldsymbol{x}(0)=\boldsymbol{x}^{(1)}$，$\boldsymbol{x}(1)=\boldsymbol{x}^{(2)}$ である．また，(5.6.8)とこの定義とから，任意の $t(\in[0,1])$に対して，

$$\boldsymbol{x}(t)^{\mathrm{H}}A_{\mathrm{R}}\boldsymbol{x}(t)=\boldsymbol{x}^{(1)\mathrm{H}}A_{\mathrm{R}}\boldsymbol{x}^{(1)}=\boldsymbol{x}^{(2)\mathrm{H}}A_{\mathrm{R}}\boldsymbol{x}^{(2)}, \quad \boldsymbol{x}(t)^{\mathrm{H}}\boldsymbol{x}(t)=1 \tag{5.6.12}$$

であることが確かめられる．したがって，$\{z(t)=\boldsymbol{x}(t)^{\mathrm{H}}A\boldsymbol{x}(t)\,|\,0\leqq t\leqq 1\}$ は $z^{(1)}$ と $z^{(2)}$ を含む虚軸に平行な直線分であることがわかる．(以上の議論だけでは，“$F(A)$がそのような直線分を含む”ということしかいえないが，$F(A)$の凸性の証明のためには，それでも十分である．) □

例 5.6.3 2×2 型の行列 A の値域 $F(A)$は，A の二つの固有値を焦点とする楕円である．実際，V_n の正規直交基底を適当に選んで A を Schur 形 $\begin{bmatrix}\lambda_1 & c\\ 0 & \lambda_2\end{bmatrix}$ に表しておく．$A'=A-\dfrac{\lambda_1+\lambda_2}{2}I$ の値域 $F(A')$は $F(A)$を$(\lambda_1+\lambda_2)/2$ だけ平行移動したものであるから，$\mathrm{tr}\,A'=0$ なる $F(A')$を考えれば十分である．さらに，(5.6.5)の形の回転によって，二つの固有値を実軸上に移すことができる．このようにして，$A=\begin{bmatrix}\alpha & 2b\\ 0 & -\alpha\end{bmatrix}(\alpha\geqq 0)$という形の行列について考えれば十分であることがわかる．さらに，$b=\beta\mathrm{e}^{\mathrm{i}\gamma}(\beta\geqq 0$，$\gamma$は実数)としたとき，$Q=\begin{bmatrix}\mathrm{e}^{\mathrm{i}\gamma} & 0\\ 0 & 1\end{bmatrix}$なるユニタリ行列を用いて A を $Q^{\mathrm{H}}AQ$ と変換しておけば $B=\begin{bmatrix}\alpha & 2\beta\\ 0 & -\alpha\end{bmatrix}(\alpha,\beta\geqq 0)$という形の行列を考えれば十分であることにもなる．さて，この形の行列を(5.6.6)

のように分解してみる：

$$B=B_{\mathrm{R}}+\mathrm{i}B_{\mathrm{I}},\quad B_{\mathrm{R}}=\begin{bmatrix}\alpha & \beta\\ \beta & -\alpha\end{bmatrix},\quad B_{\mathrm{I}}=\begin{bmatrix}0 & -\mathrm{i}\beta\\ \mathrm{i}\beta & 0\end{bmatrix}. \tag{5.6.13}$$

$B(\theta)=\mathrm{e}^{-\mathrm{i}\theta}B$ に対して，$F(B(\theta))$ は $F(B)$ を原点のまわりに $-\theta$ だけ回転した図形になるが，この“横幅”は Hermite 行列

$$B_{\mathrm{R}}(\theta)=B_{\mathrm{R}}\cos\theta+B_{\mathrm{I}}\sin\theta=\begin{bmatrix}\alpha\cos\theta & \beta\mathrm{e}^{-\mathrm{i}\theta}\\ \beta\mathrm{e}^{\mathrm{i}\theta} & -\alpha\cos\theta\end{bmatrix} \tag{5.6.14}$$

の固有値($\mathrm{tr}\,B_{\mathrm{R}}(\theta)=0$ であるから，絶対値が等しく符号が反対)で定められる．二つの固有値はもちろん

$$\lambda_{1,2}(\theta)=\pm\sqrt{\alpha^2\cos^2\theta+\beta^2}=\pm\sqrt{(\alpha^2+\beta^2)\cos^2\theta+\beta^2\sin^2\theta} \tag{5.6.15}$$

である．ところで，初等的な計算でわかるように，$(-\alpha,0)$，$(\alpha,0)$ を焦点とし β を短半径とする楕円($a=\sqrt{\alpha^2+\beta^2}$ が長半径，$e=\alpha/\sqrt{\alpha^2+\beta^2}$ が離心率)の，法線の方向余弦が $(\cos\theta,\sin\theta)$ であるような接線までの原点からの距離は

$$l(\theta)=\sqrt{a^2\cos^2\theta+\beta^2\sin^2\theta} \tag{5.6.16}$$

である．逆に，関係(5.6.16)を満たすような凸図形が楕円であることも知られている．(5.6.15)と(5.6.16)を比較することによって，$F(B)$ が α，$-\alpha$ を焦点とする離心率 $\alpha/\sqrt{\alpha^2+\beta^2}$ の楕円であることが示される(図 5.6.1 も参照)．　□

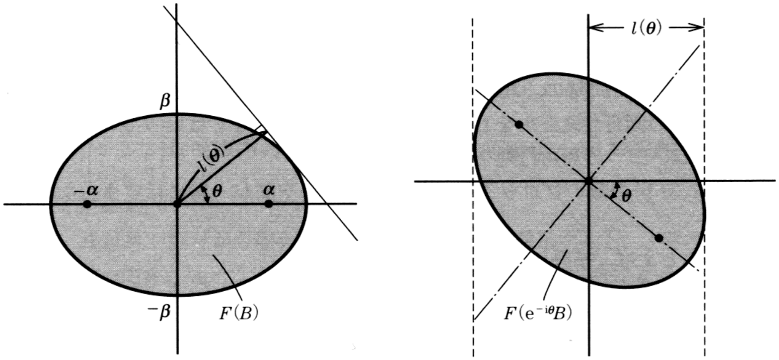


図 5.6.1

演習問題

5.1　巡回行列の固有値と固有ベクトルを求めよ．

［略解］　2.16.7項の(2.16.61)は例2.14.7の(2.14.36)で定義される Z の第 k 列ベクトルが巡回行列 C の固有値 $\sqrt{n}\gamma_{k-1}$((2.16.59)で定義される)に対応する固有ベクトルであることを示している．

5.2　二つの n 次正方行列 A, B について，次を示せ．

(i)　AB の固有値と BA の固有値は，多重度も含めて一致する．

(ii)　A と B が交換可能，すなわち $AB=BA$ を満たすとき，A と B は共通の固有ベクトルをもつ．

［略解］　(i)　AB と BA が同じ特性多項式をもつ，すなわち $\varphi_{AB}(\lambda)=\varphi_{BA}(\lambda)$ であることを示せばよい．ところで，$\varphi_{AB}(\lambda)$ の $(-\lambda)^{n-k}$ の係数は注意5.1.5で述べたように AB の k 次の主小行列式の和である．それは **Binet-Cauchy の公式**(2.8.3項，特に(2.8.17))と(2.7.3)により $\mathrm{tr}(AB)^{<k>}=\mathrm{tr}(A^{<k>}B^{<k>})=\mathrm{tr}(B^{<k>}A^{<k>})=\mathrm{tr}(BA)^{<k>}$ であるから，φ_{BA} の $(-\lambda)^{n-k}$ の係数に等しい．　(ii)　A の一つの固有ベクトルを $\boldsymbol{y}(\neq\boldsymbol{0})$ とすると，$A\boldsymbol{y}=\lambda\boldsymbol{y}$．$AB=BA$ により，$A(B\boldsymbol{y})=BA\boldsymbol{y}=\lambda B\boldsymbol{y}$．$A(B^k\boldsymbol{y})=B(AB^{k-1})\boldsymbol{y}=\cdots=B^kA\boldsymbol{y}=\lambda(B^k\boldsymbol{y})$ により，$B\boldsymbol{y}$，$B^2\boldsymbol{y}$，…も――零ベクトルにならない限り―― A の固有値 λ に対応する固有ベクトルである．$\{\boldsymbol{y}, B\boldsymbol{y}, B^2\boldsymbol{y}, \cdots, B^{n-1}\boldsymbol{y}\}$ が張る V_n の部分空間 $L\equiv L(\boldsymbol{y}, B\boldsymbol{y}, \cdots, B^{n-1}\boldsymbol{y})$ (Cayley-Hamilton の定理により $B^k\boldsymbol{y}$ $(k\geqq n)$ は $\{B^k\boldsymbol{y}\mid k\leqq n-1\}$ に従属)の任意のベクトルも A の固有値 λ に対応する固有ベクトルである．また，L は B の不変部分空間でもある．すなわち $\forall\boldsymbol{z}\in L$ に対して $B\boldsymbol{z}\in L$．L に適当に基底 $\{\boldsymbol{e}_1, \cdots, \boldsymbol{e}_m\}$ $(\boldsymbol{e}_i\in L)$ を選び，$B: L\to L$ の $(m\times m$ 型の$)$ 表現行列 $\tilde{B}$ $(m\leqq n)$ の固有ベクトルを $\boldsymbol{f}$ とすれば，$\tilde{B}\boldsymbol{f}=\mu\boldsymbol{f}$．$L$ から V_n に戻って考えれば，$\boldsymbol{x}=\sum_{i=1}^{m}f_i\boldsymbol{e}_i$ は $B\boldsymbol{x}=\mu\boldsymbol{x}$ を満たすと同時に，$A\boldsymbol{e}_i=\lambda\boldsymbol{e}_i$ であるから，$A\boldsymbol{x}=\lambda\boldsymbol{x}$ をも満たす．すなわち，A と B は共通の固有ベクトルをもつ．固有値まで一致するとはいっていないことに注意．(たとえば，二つの対角行列を考えてみよ．)

5.3　(5.2.31 b)の形のブロック化に対応する部分空間が一意には定まらないことを例で示せ．

［略解］　たとえば，基底 $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3, \boldsymbol{e}_4, \boldsymbol{e}_5\}$ に関して

$$\left[\begin{array}{ccc|cc} 0 & 1 & 0 & & \\ 0 & 0 & 1 & \multicolumn{2}{c}{O} \\ 0 & 0 & 0 & & \\ \hline & O & & 0 & 1 \\ & & & 0 & 0 \end{array}\right]$$

という行列があるとき，基底 $\{\boldsymbol{e}_1+\boldsymbol{e}_4, \boldsymbol{e}_2+\boldsymbol{e}_5, \boldsymbol{e}_3, \boldsymbol{e}_4, \boldsymbol{e}_5\}$ に関する行列表現を作ってみよ．

5.4　Gershgorin の定理(強い形)において，A の Gershgorin の円 $C_1, \cdots, C_n$ の合併

集合の一つの連結成分が2個以上の円から構成されていてかつそれらの中に固有値を含まない円があるような例を作れ．

［略解］ $A=\begin{bmatrix} 0 & 3 \\ -7 & 8 \end{bmatrix}$について計算してみよ．さらに $A(t)=\begin{bmatrix} 0 & 3t \\ -7t & 8 \end{bmatrix}$について，$t=0\sim1$ のときの固有値 $\lambda_1(t), \lambda_2(t)$ と Gershgorin の円 $C_1(t), C_2(t)$ の動きを調べよ．特に，$t=0.7, 0.8, 0.85, 0.9$ に対して調べよ．

5.5 z の n 次多項式 $p(z)=z^n+a_{n-1}z^{n-1}+\cdots+a_1z+a_0$ $(a_i\in\mathbb{C})$ と相異なる複素数 $z_1,\cdots,z_n$ に対して，$\varphi_i=-p(z_i)/\prod_{j\neq i}(z_i-z_j)$ とおく．このとき，

$$A=\begin{bmatrix} z_1+\varphi_1 & \varphi_1 & \varphi_1 & \cdots & \varphi_1 \\ \varphi_2 & z_2+\varphi_2 & \varphi_2 & \cdots & \varphi_2 \\ \varphi_3 & \varphi_3 & z_3+\varphi_3 & \cdots & \varphi_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \varphi_n & \varphi_n & \varphi_n & \cdots & z_n+\varphi_n \end{bmatrix}$$

の特性多項式 $\varphi_A(\lambda)$ は $(-1)^n p(\lambda)$ に等しいことを示せ．このことから，$p(z)$ が n 個の零点 $\alpha_1,\cdots,\alpha_n$ をもち，$z_1,\cdots,z_n$ がそれらの近似値とみなせるくらい各 Gershgorin の円が小さいとする．このとき

$$|z_i+\varphi_i-\alpha_j|\leqq(n-1)|\varphi_i|$$

により $z_i+\varphi_i$ と α_j の差が見積れることも示せ．

［略解］ $\varphi_A(\lambda)$ は λ の n 次多項式で λ^n の係数は $(-1)^n$．また $\varphi_A(\lambda)$ を $z_1-\lambda,\cdots,z_n-\lambda$ の多項式の形に展開すると，$\varphi_A(\lambda)=\prod_{i=1}^{n}(z_i-\lambda)+\sum_{i=1}^{n}\left(\prod_{j\neq i}(z_i-\lambda)\right)\varphi_i=\prod_{i=1}^{n}(z_i-\lambda)-\sum_{i=1}^{n}\left(\prod_{j\neq i}(z_i-\lambda)\Big/(z_i-z_j)\right)p(z_i)$ と表せる（$(z_i-\lambda)$ の $n-2$ 個以下の積の係数は A の形からして0になることに注意！）ので，$\varphi_A(z_k)=(-1)^n p(z_k)$ $(k=1,\cdots,n)$ である．ゆえに，$\varphi_A(\lambda)=(-1)^n p(\lambda)$．このことは $p(z)$ の零点が A の固有値に等しいことを示している．Gershgorin の定理を適用すると，所望の結果が得られる．

5.6 $A=\begin{bmatrix} 1 & \varepsilon & 0 \\ \varepsilon & 2 & \varepsilon \\ 0 & \varepsilon & 3 \end{bmatrix}$ $(\varepsilon=10^{-4})$ の固有値の存在範囲を Gershgorin の定理を用いてなるべく詳しく定めよ．

［略解］ A は実対称だから固有値は実数．定理を直接適用すると，3個の円は交わらないから，$|\lambda_1-1|\leqq\varepsilon, |\lambda_2-2|\leqq2\varepsilon, |\lambda_3-3|\leqq\varepsilon$．$S=\mathrm{diag}[0.25,\varepsilon,\varepsilon]$ により相似変換すると $S^{-1}AS=\begin{bmatrix} 1 & 4\varepsilon^2 & 0 \\ 0.25 & 2 & \varepsilon \\ 0 & \varepsilon & 3 \end{bmatrix}$ となるので，同様にして，$|\lambda_1-1|\leqq4\varepsilon^2, |\lambda_2-2|\leqq0.25+\varepsilon$，$|\lambda_3-3|\leqq\varepsilon$（3個の円はやはり交らない）．$S=\mathrm{diag}[\varepsilon,0.5,\varepsilon]$ とおけば，$S^{-1}AS=$

$\begin{bmatrix} 1 & 0.5 & 0 \\ 2\varepsilon^2 & 2 & 2\varepsilon^2 \\ 0 & 0.5 & 3 \end{bmatrix}$ で，$|\lambda_2-2|\leqq 4\varepsilon^2$，など．$S=\mathrm{diag}[\varepsilon,\varepsilon,0.25]$ とおけば，$S^{-1}AS=$ $\begin{bmatrix} 1 & \varepsilon & 0 \\ \varepsilon & 2 & 0.25 \\ 0 & 4\varepsilon^2 & 3 \end{bmatrix}$ で，$|\lambda_3-3|\leqq 4\varepsilon^2$，など．結局，このような簡単な相似変換(スケール変換)だけで，$|\lambda_1-1|\leqq 4\varepsilon^2$，$|\lambda_2-2|\leqq 4\varepsilon^2$，$|\lambda_3-3|\leqq 4\varepsilon^2$ $(4\varepsilon^2=4\times 10^{-8})$ というずっと精度の高い固有値の近似が得られる．

5.7　A が n 次 Hermite 行列のとき，A と

$$Q=(I_n-\mathrm{i}A)(I_n+\mathrm{i}A)^{-1},\quad A=\mathrm{i}(Q+I_n)^{-1}(Q-I_n)$$

で結ばれる Q はユニタリ行列で固有値 -1 をもたないことを示せ．逆に，固有値 -1 をもたない n 次ユニタリ行列 Q に対して，上の関係で結ばれる A は Hermite 行列であることを示せ．(このような A と Q の間の変換を **Cayley 変換**という．)

［略証］ Hermite 行列 A の固有値 λ は実数であるから，$I_n+\mathrm{i}A$ の固有値 $1+\mathrm{i}\lambda$ は 0 にならない．したがって $(I_n+\mathrm{i}A)^{-1}$ が存在する．また，$A, I_n\pm\mathrm{i}A$，$(I_n\pm\mathrm{i}A)^{-1}$ などは交換可能であるから，$Q^{\mathrm{H}}Q=(I_n+\mathrm{i}A)^{-1\mathrm{H}}(I_n-\mathrm{i}A)^{\mathrm{H}}(I_n-\mathrm{i}A)(I_n+\mathrm{i}A)^{-1}=(I_n-\mathrm{i}A)^{-1}(I_n+\mathrm{i}A)(I_n-\mathrm{i}A)(I_n+\mathrm{i}A)^{-1}=I_n$．この Q は A と交換可能であるから，演習問題 5.2(ii) により，Q の固有値は $\mu=(1-\mathrm{i}\lambda)/(1+\mathrm{i}\lambda)$ (λ は A の固有値) という形をしており，$-\infty<\lambda<\infty$ のとき $(1-\mathrm{i}\lambda)/(1+\mathrm{i}\lambda)=\mathrm{e}^{\mathrm{i}\theta}$ $(-\pi<\theta<\pi)$ なので，$\mu=-1$ とはならない．逆に，Q が固有値 -1 をもたないユニタリ行列なら，$Q+I_n$ は正則で，$A=\mathrm{i}(Q+I_n)^{-1}(Q-I_n)$ が定義される．そして，

$$A^{\mathrm{H}}=-\mathrm{i}(Q^{\mathrm{H}}-I_n)(Q^{\mathrm{H}}+I_n)^{-1}=-\mathrm{i}(I_n-Q)Q^{\mathrm{H}}Q(I_n+Q)^{-1}=\mathrm{i}(Q-I_n)(Q+I_n)^{-1}=A.$$

5.8　演習問題 5.7 の Cayley 変換は実数要素の行列の場合にはどのようになるか．

［略解］ 反対称行列 A と固有値 -1 をもたない直交行列 Q は，$Q=(I-A)(I+A)^{-1}$，$A=(I+Q)^{-1}(I-Q)$ という関係により一対一に対応する．

5.9　実数データ $\lambda_i (i=1,\cdots,n), \mu$ が与えられている(ただし，$\lambda_1\geqq\lambda_2\geqq\cdots\geqq\lambda_n$，$\lambda_1\geqq\mu\geqq\lambda_n$ とする)．

$$\text{(i)}\quad p_i\geqq 0\ (i=1,\cdots,n),\quad \text{(ii)}\quad \sum_{i=1}^{n}p_i=1,\quad \text{(iii)}\quad \sum_{i=1}^{n}\lambda_i p_i=\mu$$

の条件を満たす p_i をいろいろ変化させたとき，$\sigma=\sqrt{\sum_{i=1}^{n}\lambda_i^2 p_i-\mu^2}$ の取りうる値の最大値 $\sigma_{\max}$ と最小値 $\sigma_{\min}$ を求めよ．

［略解］ これは，λ_i，$\cdots$，λ_n，μ が与えられたとき (i), (ii), (iii) を制約条件とし $\sigma^2\left(=\sum_{i=1}^{n}\lambda_i^2 p_i-\mu^2\right)$ を目的関数とする，変数 p_i に関する**線形計画問題**(linear programming problem)である．したがって，(ii)，(iii) を満たす**基底解**(→4.3.3 項)の中から (i) を満たす最適解を捜せばよい(線形計画法の“基底定理”)．$\lambda_1=\lambda_2=\cdots=\lambda_n$ の場合は

トリビアルであるから，$\lambda_i \neq \lambda_j$ なる i,j が存在する場合を考える．基底解は，$\lambda_i \neq \lambda_j$ を満たす($\lambda_i > \lambda_j$ とする) i,j に対して，(ii)，(iii)を p_i, p_j について解いた式

$$(*1)\quad p_i + \frac{1}{\lambda_i - \lambda_j} \sum_{k \neq i,j} (\lambda_k - \lambda_j) p_k = \frac{\mu - \lambda_j}{\lambda_i - \lambda_j},$$

$$(*2)\quad p_j + \frac{1}{\lambda_i - \lambda_j} \sum_{k \neq i,j} (\lambda_i - \lambda_k) p_k = \frac{\lambda_i - \mu}{\lambda_i - \lambda_j}$$

において，$p_k = 0$ $(k \neq i,j)$ とおいたものである．このとき $p_i = (\mu - \lambda_j)/(\lambda_i - \lambda_j)$，$p_j = (\lambda_i - \mu)/(\lambda_i - \lambda_j)$ であるから，この基底解が実行可能，すなわち $p_i \geqq 0$，$p_j \geqq 0$ であるためには，以下の(∗3)が成り立たなければならない：

$$(*3)\quad \lambda_i \geqq \mu \geqq \lambda_j.$$

次に，σ^2 の表式に(∗1)，(∗2)を代入して p_i, p_j を消去すると，

$$\begin{aligned}(*4)\quad \sigma^2 &= -\mu^2 + \frac{(\mu - \lambda_j)\lambda_i{}^2 + (\lambda_i - \mu)\lambda_j{}^2}{\lambda_i - \lambda_j} \\ &\quad + \sum_{k \neq i,j} \left(\lambda_k{}^2 - \frac{(\lambda_k - \lambda_j)\lambda_i{}^2 + (\lambda_i - \lambda_k)\lambda_j{}^2}{\lambda_i - \lambda_j} \right) p_k \\ &= -\mu^2 + \mu(\lambda_i + \lambda_j) - \lambda_i\lambda_j + \sum_{k \neq i,j} (\lambda_k{}^2 - \lambda_k(\lambda_i + \lambda_j) + \lambda_i\lambda_j) p_k \\ &= (\lambda_i - \mu)(\mu - \lambda_j) - \sum_{k \neq i,j} (\lambda_i - \lambda_k)(\lambda_k - \lambda_j) p_k\end{aligned}$$

が得られる．上の実行可能基底解において σ^2 が最大(最小)であるためには，すべての $k(\neq i,j)$ に対して $(\lambda_i - \lambda_k)(\lambda_k - \lambda_j) \geqq 0\,(\leqq 0)$ でなければならない．$\lambda_i > \lambda_j$ を考慮すると，$\lambda_i \geqq \lambda_k \geqq \lambda_j$ ($\lambda_k \geqq \lambda_i$ あるいは $\lambda_k \leqq \lambda_j$)がすべての $k(\neq i,j)$ に対して成り立たなければならない．それには，$\lambda_i = \lambda_1, \lambda_j = \lambda_n$ (λ_i と λ_j は μ を挟む隣り合う λ_l)でなければならない．そしてこのとき，$\sigma^2 = (\lambda_i - \mu)(\mu - \lambda_j)$．ゆえに，

$$(*5)\quad \sigma_{\max} = \sqrt{(\lambda_1 - \mu)(\mu - \lambda_n)},$$

$$(*6)\quad \sigma_{\min} = \sqrt{(\lambda_i - \mu)(\mu - \lambda_{i+1})} \qquad (\lambda_i \geqq \mu \geqq \lambda_{i+1})$$

である．

5.10　n 次正方行列 $A = [a^i_j]$，単位行列 $I_n = [\delta^i_j]$ に対して，

$$(*1)\quad (n+1) \sum_{i_1,\cdots,i_n=1,\cdots,n} a^{i_1}_{[i_1} a^{i_2}_{i_2} \cdots a^{i_n}_{i_n} \delta^i_{j]} \qquad (i,j=1,\cdots,n)$$

は，値 $1,\cdots,n$ を取る $n+1$ 個の添字 $i_1,\cdots,i_n,j$ に関する交代化演算を含んでいるから，恒等的に 0 である．この式を適当に変形すると，Cayley-Hamilton の定理が得られることを示せ．[ヒント：2.5.2 項を利用せよ．$\sum_{i_1,\cdots,i_k} a^{i_1}_j a^{i_2}_{i_1} a^{i_3}_{i_2} \cdots a^{i_k}_{i_{k-1}} \delta^i_{i_k}$ は A^k の(i,j)要素であり，$\sum_{i_1,\cdots,i_k} a^{i_1}_{[i_1} a^{i_2}_{i_2} \cdots a^{i_k}_{i_k]} = c_{n-k}$ であることにも注意．]

［略解］　式(∗1)の下側の交代化演算において$(n+1)$個の $i_1, i_2, \cdots, i_n, j$ を$(n+1)$個の点①,⋯,ⓝ,⓪にそれぞれ対応させる．$(n+1)![i_1, i_2, \cdots, i_n, j]$の各順列を，これらの点を結ぶサイクルで表現する(→ 補章 A.3 節なども参照)．辺ⓚ→ⓛの重みは $a^{i_k}_{i_l}$，⓪→ⓚの重みは $a^i_{i_k}$，ⓚ→⓪の重みは $a^{i_k}_j$ であるとする．⓪↻の重みは δ^i_j とする．サイクルの

重みはそれを構成する辺の重みの積とする．すると，

〈a〉 ⓪を含む長さが $r+1$ のサイクルは，全部で ${}_nP_r$ 個あるが，それぞれの重みは $a^{i}_{i_1}a^{i_1}_{i_2}\cdots a^{i_r}_{j}$ で，$i_1,\cdots,i_r(=1,\cdots,n)$ についての和をとると $(A^r)^i_j$ に等しい．

〈b〉 残りの部分の⓪を含まないサイクルは $\sum\limits_{\substack{i_{r+1},\cdots,i_n\\=1,2,\cdots,n}} a^{[i_{r+1}}_{i_{r+1}}\cdots a^{i_n]}_{i_n}$ という形に対応するが，これは注意 2.7.1 にも指摘してあるように A の $(n-r)$ 次の主小行列式の和，すなわち A の特性方程式 $\varphi_A(x)$ の x^r の係数に等しい．

〈c〉 結果として同一の値の重みをもつサイクルの個数は，〈a〉で ${}_nP_r=(n-r)!$ 個あるが〈b〉の $[i_{r+1},\cdots,i_n]$ は $(n-r)!$ 倍してはじめて $(n-r)$ 個の記号のすべての順列となるので，結局どの r に対しても n 個の記号 $i_1,\cdots,i_n$ のすべての順列を与えることになる．式(∗1)は $(n+1)$ 個の記号のすべての順列を作ることを目標にしているので $(n+1)!$ を掛けた式になるはずであるが，その中の $n!$ はこのようにして自然に重複して数えられているので，$n+1$ だけが係数として残るわけである．（もっとも，$=0$ であるから，係数はどうでもよいが．）

このようにして純粋に組合せ論的な式(∗1)が Cayley-Hamilton の定理に他ならないことが判明する．

5.11　単因子，最小多項式を用いて，

(i)　べき等な n 次正方行列 A（$A^2=A$ を満たす行列），

(ii)　べき零な n 次正方行列 A（ある自然数 p に対して $A^p=O, A^{p-1}\neq O$ を満たす行列）

の構造を調べよ．

［略解］　(i)　$A^2-A=O$ であるから A の最小多項式は $f(x)=x^2-x$ の因子である．$\psi_A(x)=x$ なら $A=O$，$\psi_A(x)=x-1$ なら $A=I$．$\psi_A(x)=x(x-1)$ なら A のすべての単因子は $x(x-1)$，x，$x-1$，1 のいずれかである（並ぶ順序は，$(\overbrace{1,\cdots,1}^{n_1}, \overbrace{x,\cdots,x}^{n_2}, \overbrace{x(x-1),\cdots,x(x-1)}^{n_3})$，あるいは $(\overbrace{1,\cdots,1}^{n_1}, \overbrace{x-1,\cdots,x-1}^{n_2}, \overbrace{x(x-1),\cdots,x(x-1)}^{n_3})$ で，$n_1=n_3$ である（$n_1+n_2+n_3=n, n_2+2n_3=n$ でなければならないから）；$n_2=0$ であってもよい）．したがって，A は(5.4.35)，(5.4.36)において $C_i=[0],[1],\begin{bmatrix}0&0\\1&1\end{bmatrix}$ の形の対角ブロックからなるブロック対角行列に相似である．$S=\begin{bmatrix}1&0\\-1&1\end{bmatrix}$，$S^{-1}=\begin{bmatrix}1&0\\1&1\end{bmatrix}$ に対して $S^{-1}\begin{bmatrix}0&0\\1&1\end{bmatrix}S=\begin{bmatrix}0&0\\0&1\end{bmatrix}$ であるから，要するに，A は $\mathrm{diag}[0,\cdots,0,1,\cdots,1]$ に相似である．$A^2=A$ は**射影**の条件（→ 3.7.1 項，(3.7.3)）でもあることに注意．　(ii)　A の最小多項式は $\psi_A(x)=x^p$ である．したがって，単因子は $x^{p_1},x^{p_2},\cdots,x^{p_n}$（$p_1\leqq p_2\leqq\cdots\leqq p_{n-1}\leqq p_n=p$，$\sum p_i=n$）である．最小多項式（＝特性多項式）$x^{p_i}$ をもつコンパニオン行列は

$\begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & \ddots & & & \vdots \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ O & & & 1 & 0 \end{bmatrix} p_i$ であるから，A はこのような対角ブロックからなるブロック対角行列に相似である．

5.12　$A=BC$ のとき $\|A\|_F \leqq \|B\|_2 \cdot \|C\|_F$，$\|A\|_F \leqq \|B\|_F \cdot \|C\|_2$ であることを示せ．

［略証］　A の第 j 列を $\boldsymbol{a}_{\cdot j}$, C の第 j 列を $\boldsymbol{c}_{\cdot j}$ とすれば，$\|\boldsymbol{a}_{\cdot j}\|_2 \leqq \|B\|_2 \|\boldsymbol{c}_{\cdot j}\|_2$（$\|B\|_2$ の定義）．一方，$\|A\|_F{}^2 = \sum_{j=1}^{n} \|\boldsymbol{a}_{\cdot j}\|_2{}^2$ など．第 2 の不等式については A, B の行ベクトルを考えればよい．

5.13　n 次ユニタリ行列 Q, Q_1, Q_2 に対して，$\|Q\|_F = \sqrt{n}$，$\|Q\|_2 = 1$，$\|Q_1 A Q_2\|_F = \|A\|_F$，$\|Q_1 A Q_2\|_2 = \|A\|_2$ であることを示せ．

［略解］　$\forall \boldsymbol{x} \in V_n$ に対して $\|Q\boldsymbol{x}\|_2 = \|\boldsymbol{x}\|_2$ だから $\|Q\|_2 = 1$．演習問題 5.12 により $\|QA\|_F \leqq \|Q\|_2 \|A\|_F = \|A\|_F$，$\|A\|_F = \|Q^H QA\|_F \leqq \|QA\|_F$ で，結局，$\|A\|_F = \|QA\|_F$．同様にして $\|A\|_F = \|AQ\|_F$，ゆえに $\|Q_1 A Q_2\|_F = \|A\|_F$；$\sqrt{n} = \|I\|_F = \|Q^H Q\|_F = \|Q\|_F$；$\|Q_1 A Q_2\|_2 = \|A\|_2$ は，ノルムの性質 $\|AB\| \leqq \|A\| \cdot \|B\|$ と $\|Q\|_2 = 1$ より同様にして示せる．

5.14　n 次正方行列 A に対して $\|A^H A - AA^H\|_F{}^2 = 2(\|AA^H\|_F{}^2 - \|A^2\|_F{}^2)$ であることを示せ（この量は，A が正規行列のときには 0 となるので，“正規性からのずれ”を表す一つの尺度である）．

［略解］　$\|B\|_F{}^2 = \mathrm{tr}(B^H B)$，$\mathrm{tr}(BC) = \mathrm{tr}(CB)$ という関係（(2.7.3)，(2.7.4) など）を用いて $\|A^H A - AA^H\|_F{}^2$ を変形せよ．

5.15　任意の体 K の上で，

(i)　コンパニオン行列 C は，ある対称行列 T によって，その転置 C^T に相似変換できること：$C^T = T^{-1} C T$　$(T^T = T)$ を示せ．

(ii)　任意の n 次正方行列 A は，ある対称行列 S によって，その転置 A^T に相似変換できること：$A^T = S^{-1} A S$　$(S^T = S)$ を示せ．

(iii)　任意の n 次正方行列 A は対称行列 S_1，S_2（少なくとも一方は正則）の積に表せること：$A = S_1 \cdot S_2$　$(S_i{}^T = S_i)$ を示せ．

(iv)　特に $K = \mathbb{C}$ のときには，任意の正方行列 A は相似変換によって対称行列 $B = X^{-1} A X$（今度は X は対称とは限らない正則行列）に変換できることを示せ．

［略証］

(i) $C=\begin{bmatrix} 0 & & & & f_0 \\ 1 & \ddots & & O & \vdots \\ & \ddots & \ddots & & \vdots \\ & & \ddots & \ddots & \vdots \\ & O & & \ddots \quad 0 & f_{n-2} \\ & & & 1 & f_{n-1} \end{bmatrix}$ に対して，

$$T=\begin{bmatrix} f_1 & f_2 & \cdots & \cdots & f_{n-1} & -1 \\ f_2 & & \iddots & \iddots & \iddots & \\ \vdots & \iddots & \iddots & \iddots & & \\ \vdots & \iddots & \iddots & & O & \\ f_{n-1} & \iddots & & & & \\ -1 & & & & & \end{bmatrix}$$

($f_1,\cdots,f_{n-1},-1,0,0,\cdots$が斜めに並ぶ Hankel 行列)

とおけば，

$$CT=\begin{bmatrix} -f_0 & 0 & 0 & \cdots & \cdots & 0 & 0 \\ 0 & f_2 & f_3 & \cdots & \cdots & f_{n-1} & -1 \\ 0 & f_3 & & & \iddots & \iddots & \\ \vdots & \vdots & & \iddots & \iddots & & \\ \vdots & \vdots & \iddots & \iddots & & O & \\ 0 & f_{n-1} & \iddots & & & & \\ 0 & -1 & & & & & \end{bmatrix}$$

となる．T も CT も対称であり，$\det T=\pm 1$ であるから，$CT=(CT)^{\top}=TC^{\top}$，すなわち $C^{\top}=T^{-1}CT$．

(ii) A を相似変換により $X^{-1}AX=C=C_1\dot{+}\cdots\dot{+}C_s$ (C_i はコンパニオン行列)という形に変換し，各 C_i を $C_i^{\top}$ に変換する対称行列 T_i の直和を $T=T_1\dot{+}\cdots\dot{+}T_s$ とすれば，$C^{\top}=T^{-1}CT$．$S=XTX^{\top}$ とおくと，$S^{-1}AS=(X^{\top})^{-1}T^{-1}X^{-1}AXTX^{\top}=(X^{\top})^{-1}C^{\top}X^{\top}=(XCX^{-1})^{\top}=A^{\top}$，$S^{\top}=XT^{\top}X^{\top}=S$．

(iii) (ii)において $A^{\top}=S^{-1}AS$ を $SA^{\top}=AS$ と書いてみると，$(SA^{\top})^{\top}=(AS)^{\top}=SA^{\top}$ であるから，$S_1=SA^{\top}$ は対称．そこで $S_2=S^{-1}$ (これも対称で，かつ正則)とおけば，$A=SA^{\top}S^{-1}=S_1S_2$．

(iv) (i)により $A^{\top}=S^{-1}AS$ を満たす $S(=S^{\top})$を求め，S を(2.15.59)の形に Cholesky 分解(→ 2.15.4 項 a)したものを $S=YDY^{\top}$ とすると(ここまでは有理演算で可能)，S が正則であるから $D=\mathrm{diag}[d_1,\cdots,d_n]$，$d_i\neq 0$ である；各 d_i の平方根 $\sqrt{d_i}$ を用いて(ここで $K=\mathbb{C}$ であることが効いてくる；$d\in K$ に対して $\sqrt{d}\in K$ であるような体 K なら特に $K=\mathbb{C}$ でなくてもよい) $X=Y\,\mathrm{diag}[\sqrt{d_1},\cdots,\sqrt{d_n}]$ とおけば，$S=$

$XX^\top$（X は正則）．$B=X^{-1}AX$ とおけば，$B^\top=X^\top A^\top(X^{-1})^\top=X^\top S^{-1}AS(X^{-1})^\top$．$S=XX^\top$ だから $S(X^{-1})^\top=S(X^\top)^{-1}=X$，$X^\top S^{-1}=X^{-1}$ で，$B^\top=X^{-1}AX=B$，すなわち B は対称．

5.16　n 次正方行列 A が $A=A_{11}\oplus A_{22}$ と直和に分解されるとき，A の値域 $F(A)$ は A_{11} の値域 $F(A_{11})$ と A_{22} の値域 $F(A_{22})$ を両方とも含む最小の凸集合（$F(A_{11})$ と $F(A_{22})$ の凸包）であることを示せ．

［略解］　$\boldsymbol{x}=\boldsymbol{x}_1+\boldsymbol{x}_2$（$\boldsymbol{x}_i$ は不変部分空間のベクトル）と表しておくと，$\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}=\boldsymbol{x}_1{}^{\mathrm{H}}A_{11}\boldsymbol{x}_1+\boldsymbol{x}_2{}^{\mathrm{H}}A_{22}\boldsymbol{x}_2$，$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=\boldsymbol{x}_1{}^{\mathrm{H}}\boldsymbol{x}_1+\boldsymbol{x}_2{}^{\mathrm{H}}\boldsymbol{x}_2$ より，$\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}/\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=(\boldsymbol{x}_1{}^{\mathrm{H}}A_{11}\boldsymbol{x}_1/\boldsymbol{x}_1{}^{\mathrm{H}}\boldsymbol{x}_1)(\boldsymbol{x}_1{}^{\mathrm{H}}\boldsymbol{x}_1/\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x})+(\boldsymbol{x}_2{}^{\mathrm{H}}A_{22}\boldsymbol{x}_2/\boldsymbol{x}_2{}^{\mathrm{H}}\boldsymbol{x}_2)(\boldsymbol{x}_2{}^{\mathrm{H}}\boldsymbol{x}_2/\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x})$，$\boldsymbol{x}_1{}^{\mathrm{H}}\boldsymbol{x}_1/\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}+\boldsymbol{x}_2{}^{\mathrm{H}}\boldsymbol{x}_2/\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}=1$．

5.17　n 次正方行列 A の固有値が $\lambda_1,\cdots,\lambda_n$ であるとき $A^{<r>}$ は $\binom{n}{r}$ 次の正方行列であるが(→ 2.7.1 項)，$A^{<r>}$ の固有値は $\binom{n}{r}$ 個の積 $\lambda_{i_1}\cdots\lambda_{i_r}$($i_1,\cdots,i_r$ は互いに異なる)に等しいことを示せ．特に $\det A^{<r>}=(\det A)^{\binom{n-1}{r-1}}$ となることを示せ．

［略解］　A を Schur 形に変換する行列を S とすると，$SAS^{-1}=\begin{bmatrix}\lambda_1 & & * \\ & \ddots & \\ O & & \lambda_n\end{bmatrix}$であるが，

(2.8.18)により $S^{<r>}A^{<r>}(S^{-1})^{<r>}=\begin{bmatrix}\lambda_1 & & * \\ & \ddots & \\ O & & \lambda_n\end{bmatrix}^{<r>}$ であるので，$A^{<r>}$ の固有値が $\lambda_{i_1}\cdots\lambda_{i_r}$ の形になることがわかる．行列式は固有値の積に等しいから後半も示される．(固有値を経由しないでも後半の証明は可能である．)

5.18　$K=\mathbb{C}$ 上の n 次正方行列 $A=[a_{ij}]$ に対して $A=B+\mathrm{i}C$（$B^{\mathrm{H}}=B$，$C^{\mathrm{H}}=C$）と分解し，$a=\max_{i,j}|a_{ij}|$，$b=\max_{i,j}|b_{ij}|$，$c=\max_{i,j}|c_{ij}|$ とおく．このとき A の任意の固有値を $\sigma=\alpha+\mathrm{i}\beta$（$\alpha,\beta$：実数）とすると，$|\sigma|\leqq na$，$|\alpha|\leqq nb$，$|\beta|\leqq nc$ であることを示せ．

［略解］　A の固有値 σ に対応する固有ベクトルを $\boldsymbol{x}(\neq\boldsymbol{0})$ とすれば，$A\boldsymbol{x}=\sigma\boldsymbol{x}$ より $\sigma\|\boldsymbol{x}\|_2{}^2=\sum_{i,j}a_{ij}\bar{x}_ix_j$，$\bar{\sigma}\|\boldsymbol{x}\|_2{}^2=\sum_{i,j}\bar{a}_{ij}x_i\bar{x}_j$．したがって，$\alpha\|x\|_2{}^2=\sum_{i,j}b_{ij}\bar{x}_ix_j$，$\beta\|\boldsymbol{x}\|_2{}^2=\sum_{i,j}c_{ij}\bar{x}_ix_j$．そして，$|\sigma|\cdot\|\boldsymbol{x}\|_2{}^2\leqq a\sum_{i,j}|x_i|\cdot|x_j|=a(\sum_i|x_i|)^2$．Cauchy - Schwarz の不等式(→ 例 2.16.1)をベクトル$[1,\cdots,1]$，$[|x_1|,\cdots,|x_n|]$に対して適用すれば$(\sum_i|x_i|)^2\leqq n\|\boldsymbol{x}\|_2{}^2$ であるから，$|\sigma|\,\|\boldsymbol{x}\|_2{}^2\leqq na\|\boldsymbol{x}\|_2{}^2$．よって $|\sigma|\leqq na$ を得る．$|\alpha|\leqq nb$，$|\beta|\leqq nc$ についても同様．(A が Hermite 行列なら $C=O$，$c=0$ であるから $\beta=0$，すなわち σ は実数であり，A が歪 Hermite 行列なら $B=O$，$b=0$ であるから $\alpha=0$，すなわち σ は純虚数である．これはよく知られた性質の別証にもなる．)

6

行列の標準形と応用

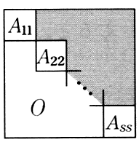
前章までに述べてきたことを，“行列の標準形”という視点からまとめ直し，いくらかの応用例を補うのが本章の目的である．

6.1 既　約　形

K は任意の体とする(体ですらなくても，第 2 章の冒頭の注意で指摘したように，以下の議論が成り立つことがある)．

6.1.1 既　約　形

n 次正方行列 A (：$V_n \to V_n$)の行と列に同時に“同じ”置換を施して，すなわち，n 次の置換行列 P によって，$P^{\mathsf{T}}AP$ をブロック上三角行列

$P^{\mathsf{T}}AP =$  (6.1.1)

に変換する(というよりは，単に“表示”を整理したものにすぎない；→注意 2.1.1)という問題を考える．(この変換は，相似変換でもあるし，合同変換でもある．置換行列 P は直交行列であるから．) もちろん，対角ブロックはなるべく小さくしたい．

$A_{11}, A_{22}, \cdots, A_{ss}$ は，V_n の A に関する不変部分空間 U_1，商空間 $W_1 = V_n/U_1$ の不変部分空間 $U_2, \cdots$，における A の表現になっていることはいうまでもない(→ 3.2.3 項，3.3.4 項)．

A を(6.1.1)の形にすれば，$A^2, A^3, \cdots$ もみな同じ形となり，それぞれの対角

ブロック $A_{kk}^{(p)}$ は $(A_{kk})^p$ となる．そこで，行列多項式 $f(A)$ も（そして収束する行列級数も）(→注意 2.2.3，2.12.1 項) $P^{\mathrm{T}}f(A)P$ が(6.1.1)の形となり，その対角ブロックは $f(A_{kk})$ に等しい．

さらに，A のべきを作るときに，A の非零要素 a_{ij} が“打ち消し合わない”ような値である場合，たとえば，(i)　$a_{ij}\geqq 0$ であるとか，(ii)　“一般の値”を取る（代数的に独立である）とかいう場合には，$I+A+\cdots+A^{n-1}$ の各対角ブロックには 0 が現れない（すべての要素が非零である）．

以上の事実は，行列 A に付随して次のように定められる**有向グラフ**(directed graph) G の言葉で述べ直すことによって，明確に理解できる．すなわち，この問題は，そのような有向グラフの強連結成分分解の問題と等価であることは明らかであろう(→ A.1 節)．

A の行集合 I，列集合 J（いまの場合 $I=J$ である）を G の**頂点集合**(vertex set)とする．$a_{ij}\neq 0$ のとき，そしてそのときに限り，頂点 i を j に結ぶ辺が G の**辺集合**(edge set) E に含まれるようにする（図 6.1.1 を見よ）．

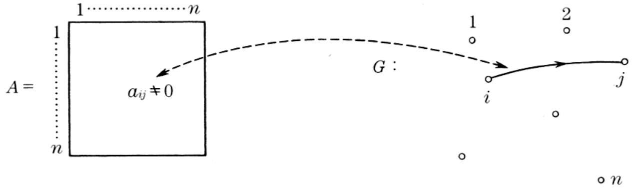


図 6.1.1　正方行列 A とそれに付随する有向グラフ G

G を**強連結成分**(strongly connected components) $G_1,\cdots,G_s$ に分解し（それに従って，頂点集合も $I_1,\cdots,I_s$ と分割される），その間の半順序に矛盾しないような線形順序に $I_1,\cdots,I_s$ を並べれば，(6.1.1)の形が得られる．したがって，(6.1.1)の形は，(i)各 I_k の中での行(列)の順序の入れ換え，(ii)半順序に矛盾しない線形順序の選び方，という自由度を除いて一意に定まる．(6.1.1)の形を，A の**既約標準形**と呼んでおく．

A^p の (i,j) 要素は，G の頂点 i から j への長さ p の**有向道**(directed path)，すなわち p 本の辺を矢印の向きに沿って辿って頂点 i から j へ行く道，が存在するときにのみ非零である（値の“打ち消し”が生じない場合には，そのとき必ず非零である）．同一の強連結成分に含まれる 2 頂点は，長さ $n-1$ 以下の（より詳

しくは I_k の中の2頂点は長さ $|I_k|-1$ 以下の）有向道で結ばれるから，$I, A, \cdots, A^{n-1}$ の (i,j) 要素は i,j が同一の連結成分に含まれるときには少なくとも一つが非零である（値の打ち消しのないとき）．（頂点 i から i へは"長さ0の"道があるとみなす．）

(6.1.1)の形において，対角ブロックがただ一つのとき，すなわち，有向グラフ G 自身が強連結であるとき，行列 A は**既約**(irreducible)であるという．既約でないとき**可約**(reducible)であるという．

注意 6.1.1　G に**有向閉路**(directed closed path)が存在しないとき（このとき G は acyclic であるといわれる），そしてそのときに限り，行列 A は行・列の順序を入れ換えることによって上三角行列にすることができる．

6.1.2　周期標準形

行列 A が既約（付随する有向グラフ G が強連結）であるとき，より詳しい"構造"が見出せる場合がある．グラフ G のすべての有向閉路の長さの最大公約数を σ とする（"すべての有向閉路"は無限個あるが，長さの最大公約数は"すべての**初等的**(elementary；「同一の頂点を2度以上通らない」の意)有向閉路"について取れば十分であるから，有限の手続きで σ を求めることができる）．

G の頂点1から長さ l $(\equiv k \bmod \sigma)$ の有向道で到達できる頂点 i の集合を I_k とする $(k=0,1,\cdots,\sigma-1)$．頂点1から頂点 i への長さ l の有向道と長さ l' の有向道があるときには $l \equiv l' \bmod \sigma$ である（実際，G の強連結性により，i から1への有向道も存在するので（図6.1.2），その一つの長さを h とすると，長さ $l+h$ と $l'+h$ の有向閉路が存在することになるが，σ の定義より $l+h=m\sigma$，$l'+h=m'\sigma$ で，$l-l'=(m-m')\sigma$ となる）から，I_k はきちんと定義される．

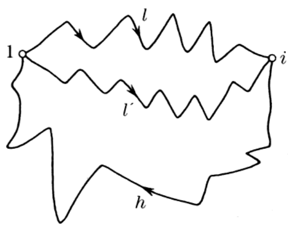


図 6.1.2

$i\in I_k$，$a_{ij}\neq 0$ なら $j\in I_{k+1}$（$k+1\geqq\sigma$ なら $k+1$ の代りに $k+1-\sigma$ とする；以下同様）であるから，行・列の集合 $I=J$ を I_0，I_1，…，$I_{\sigma-1}$ に分割して，その順序に並べ直すと，

$$
P^{\mathrm{T}}AP=
\begin{array}{c|c|c|c|c|c|c|}
 & I_0 & I_1 & I_2 & \cdots & I_{\sigma-2} & I_{\sigma-1} \\
\hline
I_0 & O & A_{01} & & \cdots & & \\
\hline
I_1 & & O & A_{12} & & & \\
\hline
I_2 & & & O & \ddots & O & \\
\hline
 & & & & \ddots & \ddots & \\
\hline
I_{\sigma-2} & & & O & & O & A_{\sigma-2,\sigma-1} \\
\hline
I_{\sigma-1} & A_{\sigma-1,0} & & & & & O \\
\hline
\end{array}
\qquad (6.1.2)
$$

という形になる．$P^{\mathrm{T}}A^2P, P^{\mathrm{T}}A^3P,\cdots$ は

$$
P^{\mathrm{T}}A^2P=
\boxed{\begin{array}{cccccc}
O & O & A_{02}^{(2)} & & & \\
\vdots & \ddots & \ddots & A_{13}^{(2)} & & O \\
\vdots & & \ddots & \ddots & \ddots & \\
O & & & \ddots & \ddots & A_{\sigma-3,\sigma-1}^{(2)} \\
A_{\sigma-2,0}^{(2)} & \ddots & & & \ddots & O \\
O & A_{\sigma-1,1}^{(2)} & O & \cdots & \cdots & O
\end{array}}
,\quad\cdots\quad,\quad\cdots \qquad (6.1.3)
$$

のようになる（一般に，$P^{\mathrm{T}}A^qP$ には $A_{k,k+q}^{(q)}=A_{k,k+1}\cdot A_{k+1,k+2}\cdot\cdots\cdot A_{k+q-1,k+q}$（行集合は I_k，列集合は I_{k+q}）という形の非零ブロックがある）．

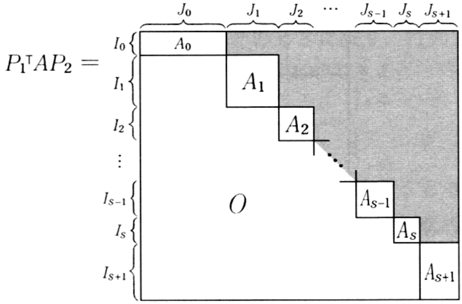
$\sigma=1$ のときには以上の構造はトリビアルになる．$\sigma=1$ のとき，A を**原始的**(primitive)と呼ぶ．一般に，σ のことを A の**周期**(period)と呼ぶ．（**原始性指数**(index of primitivity)と呼ぶ流儀もある．）(6.1.2)の形は，既約行列 A の**周期標準形**とでも呼ぶべきものであろう．

初等整数論とグラフ理論とで知られている事実「強連結グラフ G の有向閉路の長さの最大公約数が σ であれば，ある自然数 M が存在して，M より大きいすべての p（$\equiv q \mod \sigma$ とする）に対して，I_k に属する任意の頂点から I_{k+q} に属する任意の頂点への長さ p の有向道が存在する」を行列の言葉で述べ直すと，

「十分大きい p に対しては，(値の打ち消しが生じない限り) $A^{(p)}_{k,k+q}$ ($k=0,\cdots,\sigma-1$；$p\equiv q \mod \sigma$)のすべての要素が非零である」ということになる．

6.2　Dulmage-Mendelsohn 形

K は任意の体とする(体ですらなくても以下の議論が成り立つことがある)．$m\times n$ 型の行列 A (：$V_n\to U_m$)の行と列を“別々に”置換して，すなわち m 次置換行列 P_1 と n 次置換行列 P_2 を適当に選んで，



(6.2.1)

という形(**Dulmage-Mendelsohn**(ダルメージ-メンデルゾーン)**形**と呼ぶ)にする問題を考える．もちろん，このような形の中で最も“細かい”ものを求めたい．ここで，

$$|I_0|<|J_0|,\quad |I_k|=|J_k|\quad (k=1,\cdots,s),\quad |I_{s+1}|>|J_{s+1}| \tag{6.2.2}$$

である．$I_0=\emptyset, I_0=J_0=\emptyset, J_{s+1}=\emptyset, I_{s+1}=J_{s+1}=\emptyset$ あるいは $s=0$ などであってもよい(縮退した場合)．

この問題は，$A=[a_{ij}]$の行集合 I と列集合 J とを二つの頂点集合としてもち，$a_{ij}\neq 0$ のとき，そしてそのときに限り，$i\in I$ と $j\in J$ を結ぶ辺をもつ**2部グラフ**の **Dulmage-Mendelsohn 分解**の問題と等価であることは明らかであろう(伊理・藤重：「応用代数」，コロナ社，1988，1998，など参照)．

グラフ理論においては，このような分解の可能性と一意性(というより，どのくらいの自由度を残して決定されるかということ)，および分解を実行するための効率の良い算法が知られている．

注意 6.2.1　6.1 節，6.2 節で述べた事実(およびそれをさらに精密化したもの)

は，大規模システムの数値的解析の前段階において，あるいは“定性的”解析，“構造的”解析に際して，基本的な役割を果たす．

6.3 組合せ正準形

K は任意の体とする(体でなければならない)．

6.3.1 組合せ正準形

$m\times n$ 型の行列 A の**項別階数** t-rank A と**階数** rank A の間には，ほとんど自明な不等式

$$\text{t-rank}\ A \geqq \text{rank}\ A \tag{6.3.1}$$

が成り立つことは，既に 2.13.2 項で述べた(→(2.13.8))．

一方，**枢軸変換**によって A から導き出される行列(A に**組合せ論的に等価**な行列)の族 $\mathscr{P}(A)$ を考えると，他の多くの変換とは異なり，$\mathscr{P}(A)$ に属する行列の階数も項別階数も，一般には，一定でない．ところが，次の定理が知られている．

定理 6.3.1 $\mathscr{P}(A)$ の中の任意の二つの行列 B, C に対して

$$\text{t-rank}\ B \geqq \text{rank}\ C \tag{6.3.2}$$

が成り立ち，

$$\min_{B\in\mathscr{P}(A)} \text{t-rank}\ B = \max_{C\in\mathscr{P}(A)} \text{rank}\ C \quad (\equiv r_0 \text{ とおく}) \tag{6.3.3}$$

である．さらに，

$$\text{t-rank}\ A_0 = \text{rank}\ A_0 = r_0 \tag{6.3.4}$$

を満たす行列 $A_0 \in \mathscr{P}(A)$ が存在する．

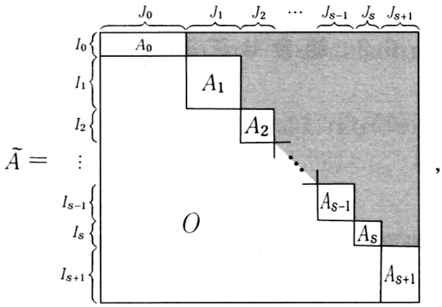
証明 脚注の文献参照[1]． ■

この定理は，マトロイド理論との関係が深く，さらにいろいろな一般化もなされているが，詳しくは脚注の文献およびそこに挙げられている諸文献を参照されたい[2)(次ページ)]．

この定理と 6.2 節の Dulmage-Mendelsohn 形とを組み合わせると，A から枢軸変換によって次の形(**組合せ正準形**(combinatorial canonical form)と呼ぶ)の

1) M. Iri : The maximum-rank minimum-term-rank theorem for the pivotal transforms of a matrix. *Linear Algebra and Its Applications*, Vol. 2 (1969), pp. 427–446.

行列 $\tilde{A}$ が得られることがわかる(述べられている諸性質を満たす最も細かい分解である). (6.2 節の縮退した場合もありうる.)



(6.3.5)

$$|I_0| < |J_0| \quad |I_k| = |J_k| \quad (k=1, \cdots, s), \quad |I_{s+1}| > |J_{s+1}|,$$
$$\text{t-rank } A_{kk} = \text{rank } A_{kk} \qquad (k=0, 1, \cdots, s, s+1),$$
$$\text{t-rank } \tilde{A} = \text{rank } \tilde{A}.$$

6.3.2 ブロック階数

行列 A の行集合 I を p 個に分割し：$I=I_1 \cup \cdots \cup I_p$ $(I_k \neq \emptyset$ $(k=1,\cdots,p)$；$I_k \cap I_l = \emptyset$ $(k \neq l))$, 列集合 J も p 個に分割し：$J=J_1 \cup \cdots \cup J_p$ $(J_k \neq \emptyset, J_k \cap J_l = \emptyset$ $(k \neq l))$, 行 I_k, 列 J_k からなる部分行列を A_k とし, $\sum_{k=1}^{p} \text{rank } A_k$ を考える. この和はもちろん I と J の分割しかたに依存するが, あらゆる分割に関してこの和の最大値を取り, それを A の **p-ブロック階数**(p-block rank)と定義し, p-rank A と記す.

明らかに

$$\text{1-rank } A = \text{rank } A \tag{6.3.6}$$

である. また A が $m \times n$ 型なら

$$p=\min(m,n) \text{ のとき } p\text{-rank } A = \text{t-rank } A \tag{6.3.7}$$

である. また,

2) M. Iri：Structural theory for the combinatorial systems characterized by submodular functions. W. R. Pulleyblank (ed.)：*Progress in Combinatorial Optimization*, Academic Press Canada, 1984, pp. 197–219.

K. Murota：*Matrices and Matroids for Systems Analysis*, Algorithms and Combinatorics, 20, Springer-Verlag, Berlin-Heidelberg, 2000.

S. Fujishige：*Submodular Functions and Optimization*, 2nd ed. Annals of Discrete Mathematics, 58, Elsevier, Amsterdam, 2005.

$$\text{rank } A = 1\text{-rank } A \leqq 2\text{-rank } A \leqq 3\text{-rank } A \leqq \cdots \leqq \text{t-rank } A \qquad (6.3.8)$$

が成り立つことも，階数の定義から，ほぼ自明であろう．

一般に，A のブロック階数は枢軸変換により変化するが，2-ブロック階数は枢軸変換により変化せず，それが(6.3.3)で定義した r_0 に等しいことが知られている(詳しくは，242 ページの脚注 1)の文献を見よ)．

6.4 階数標準形

K は任意の体とする．$m \times n$ 型行列 A の階数が r であるとき，m 次正則行列 S と n 次正則行列 T を適当に選ぶと，すなわち，適当な基本行演算と基本列演算を施すこと(同値変換)によって，

$$SAT = \begin{array}{c|c} \overbrace{\begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix}}^{r} & O \\ \hline O & O \end{array} \quad (m \times n), \qquad r = \text{rank } A \qquad (6.4.1)$$

という形に変形できることは，すでに 2.15.1 項において述べた(→注意 2.15.4)．この形を A の**階数標準形**(rank normal form)と呼んでおく(一意に定まることはいうまでもない)．

この形と，正則行列を(左および右から)掛けることによって行列の階数が変化しないという事実を利用して，やや面倒な関係式を，視覚的にわかりやすく証明し，理解することができる場合が少なくない．一例として，2.13.1 項の(2.13.6)に掲げた不等式

$$\text{rank } (AB) + \text{rank } (BC) \leqq \text{rank } (ABC) + \text{rank } B$$
$$(A : k \times l \text{ 型},\ B : l \times m \text{ 型},\ C : m \times n \text{ 型}) \qquad (6.4.2)$$

の証明をしてみよう．

まず，A, C を階数標準形に表す(S_1, T_1, S_2, T_2 は正則)．

$$S_1AT_1 = A' = \begin{array}{c|c} \begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix} & \\ \hline & O \end{array} \ \ (\text{行 } 1 \cdots r_A,\ r_A+1 \cdots k;\ \text{列 } 1 \cdots r_A,\ r_A+1 \cdots l), \qquad S_2CT_2 = C' = \begin{array}{c|c} \begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix} & \\ \hline & O \end{array} \ \ (\text{行 } 1 \cdots r_C,\ r_C+1 \cdots m;\ \text{列 } 1 \cdots r_C,\ r_C+1 \cdots n),$$

$$r_A = \text{rank } A = \text{rank } A', \qquad r_C = \text{rank } C = \text{rank } C'. \qquad (6.4.3)$$

次に

$$B' = T_1^{-1} B S_2^{-1} \tag{6.4.4}$$

とおけば，$A'B' = S_1ABS_2^{-1}$，$B'C' = T_1^{-1}BCT_2$，$A'B'C' = S_1ABCT_2$ だから，(6.4.2)に現れる行列積の階数はプライム $'$ をつけても変化しない．2.4.1項の部分行列の記法を用いると

$$A'B' = \begin{array}{c} 1 \\ \vdots \\ r_A \\ r_A+1 \\ \vdots \\ k \end{array}\overset{1 \cdots\cdots m}{\begin{bmatrix} B'^{1\cdots r_A}_{1\cdots m} \\ \hline O \end{bmatrix}}, \quad B'C' = \begin{array}{c} 1 \\ \vdots \\ l \end{array}\overset{1 \cdots r_C \ \ r_C+1 \cdots n}{\left[\begin{array}{c|c} B'^{1\cdots l}_{1\cdots r_C} & O \end{array}\right]}, \quad A'B'C' = \begin{array}{c} 1 \\ \vdots \\ r_A \\ r_A+1 \\ \vdots \\ k \end{array}\overset{1 \cdots r_C \ \ r_C+1 \cdots n}{\left[\begin{array}{c|c} B'^{1\cdots r_A}_{1\cdots r_C} & O \\ \hline O & O \end{array}\right]} \tag{6.4.5}$$

であるから，(6.4.2)を証明するには

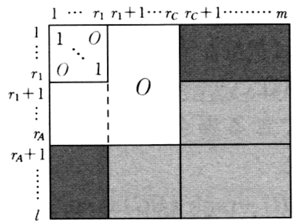
$$\mathrm{rank}\ B'^{1\cdots r_A}_{1\cdots m} + \mathrm{rank}\ B'^{1\cdots l}_{1\cdots r_C} \leqq \mathrm{rank}\ B'^{1\cdots r_A}_{1\cdots r_C} + \mathrm{rank}\ B' \tag{6.4.6}$$

を示せばよい．

$$r_1 = \mathrm{rank}\ B'^{1\cdots r_A}_{1\cdots r_C}, \quad r_2 = \mathrm{rank}\ B'^{1\cdots r_A}_{1\cdots m}, \quad r_3 = \mathrm{rank}\ B'^{1\cdots l}_{1\cdots r_C} \tag{6.4.7}$$

とおく．

B' の列 $1, \cdots, r_C$ に対する基本列演算，行 $1, \cdots, r_A$ に対する基本行演算によって B' は

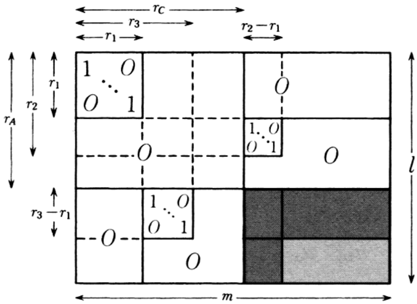
$B'' =$  (6.4.8)

という形に変形できる．この変形によって，(6.4.6)に現れる B' の部分行列(B' 自身も含めて)の階数とそれらに対応する B'' の部分行列の階数とは等しい．さらに続けて，B'' の列 $1, \cdots, r_1$ を何倍かして列 $r_C+1, \cdots, m$ に加え，行 $1, \cdots, r_1$ を何倍かして行 $r_A+1, \cdots, l$ に加えることによって(6.4.8)の ■ の部分を O にすることができる．この変形によっても問題となる部分行列の階数は変化しない．

同様にして，列 $r_1+1, \cdots, r_C$ と行 $r_A+1, \cdots, l$ に対する基本演算，行 r_1+1,

$\cdots, r_A$ と列 $r_C+1, \cdots, m$ に対する基本演算によって

$B''' =$



(6.4.9)

の形を得る．上と同様にして(6.4.9)の ■ の部分を O にすることができる．その結果得られる行列の ░ の部分の階数を r_0 とすると，

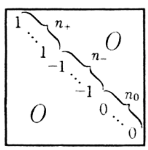
$$\operatorname{rank} B' = r_1 + (r_2 - r_1) + (r_3 - r_1) + r_0 \geqq r_2 + r_3 - r_1 \tag{6.4.10}$$

である．(6.4.10)は(6.4.6)そのものである．

6.5　Sylvester 形

$K=\mathbb{C}$ とする．n 次 Hermite 行列 A を n 次正則行列 S による合同変換によって

$S^{\mathrm{H}}AS =$



(6.5.1)

という形に変形できること，この際，**符号指数**(n_+, n_-, n_0)は S の選び方によらないこと(**Sylvester の慣性則**)などは，既に 3.5 節(特に注意 3.5.7)で述べた通りである．(6.5.1)の形も一種の標準形とみなせる．

6.6　整数行列の Hermite 標準形，Smith 標準形

$K=\mathbb{Z}$ (整数環；体ではない)とする．

6.6.1　Hermite 標準形

2.15.6 項で述べたように，整数要素の $m\times n$ 型行列 A は，適当な m 次**単模行列**(整数要素の行列で，その逆行列も整数要素からなるようなもの；行列式の値が±1 であるといっても同じ) S と n 次置換行列 P を用いて

$$\left.\begin{array}{ll} U=SAP=[u_{ij}], & \\ u_{ij}\geqq 0 & (i=1,\cdots,m\,;\,j=1,\cdots,n), \\ u_{ij}=0 & (i>j\text{ あるいは } i>r), \\ u_{ij}<u_{jj} & (i<j\,;\,j=1,\cdots,r), \\ r=\operatorname{rank} A & \end{array}\right\} \tag{6.6.1}$$

という形に変形できる．特に，$r=\operatorname{rank} A=n$ $(\leqq m)$ の場合には $P=I_n$ とすることが可能で，そうすると S も U も A により一意に定まる(→注意 2.15.19)．この形を **Hermite 標準形**と呼ぶ．

6.6.2　Smith 標準形

2.15.6 項の注意 2.15.25 で述べた通り，整数要素の $m\times n$ 型行列 A に対して，適当に m 次単模行列 S, n 次単模行列 T を選ぶと，

$$SAT=\left[\begin{array}{ccc|c} e_1 & & O & \\ & \ddots & & O \\ O & & e_r & \\ \hline & O & & O \end{array}\right] \tag{6.6.2}$$

という形にすることができる．ここで，$e_1,\cdots,e_r$ $(r=\operatorname{rank} A)$ は正の整数で e_k が e_{k+1} を割り切る $(k=1,\cdots,r-1)$．e_k は A の**単因子**と呼ばれ，A の**行列式因子** d_k (A のすべての k 次小行列式の最大公約数)を用いれば $e_k=d_k/d_{k-1}$ と表せる $(k=1,\cdots,r\,;\,d_0=1$ と約束する)．S,T の選び方によらず(6.6.2)の形は A により一意に定まる．この形を A の **Smith**(スミス)**標準形**(あるいは**単因子標準形**)と呼ぶ．同じ Smith 標準形をもつ二つの(同じ型の)行列は，左右から単模行列を掛けることによって，一方から他方に変換される．

注意 6.6.1　整数要素行列 A の Smith 標準形は，代数学(群論)における **Abel**(アーベル)**群の基本定理**，すなわち，「(有限個の生成元から生成される)Abel 群 G は，いくつかの無限位数の巡回群(整数の加群と同形)といくつかの有限位数の

巡回群の直積として一意に表すことができる．ここで，有限位数の巡回群の位数を $e_1,\cdots,e_r$ とすれば，e_k は e_{k+1} を割り切る」と同内容のものである．すなわち，群演算を加法の形に表して，G の生成元を g_1，…，g_n，その間の**基本関係**を

$$\sum_{j=1}^{n} a_{ij}g_j=0 \qquad (i=1,\cdots,m) \tag{6.6.3}$$

(0 は単位元)とすると，S は基本関係の組を等価な組に変換するものであり，T は生成元の組を等価な組に変換するものである．(6.6.2)の形に整理してしまえば，そのときの生成元 g_1'，…，g_n' $(\in G)$ に対する基本関係は

$$\left.\begin{array}{l} e_k g_k'=0 \qquad (k=1,\cdots,r), \\ g_{r+1}', \ \cdots, \ g_n' \text{ は自由元} \end{array}\right\} \tag{6.6.4}$$

となり，G は g_1'，…，g_r' が生成する巡回群(それぞれ位数 e_1，…，e_r)と，g_{r+1}'，…，g_n' が生成する $n-r$ 個の無限位数巡回群との直積で表される．

6.6.3 整数 Farkas 定理

注意 4.3.1 で述べた Farkas 型の定理の整数版が成立することを示す最もわかりやすい方法は，Smith 標準形を用いることであろう(単因子にまでもって行かなくても，単模変換で対角化できるということだけを使う)．$K=\mathbb{Q}$ (有理数体)の上で扱う．

整数 Farkas 定理　$K=\mathbb{Q}$ の上の $m\times n$ 型行列 A と n 次元縦ベクトル $\boldsymbol{c}$ が

(イ) “$A\boldsymbol{x}$ が整数を要素とするベクトルになるようなすべての n 次元ベクトル $\boldsymbol{x}$ に対して $\boldsymbol{c}^{\top}\cdot\boldsymbol{x}$ は整数である”

という性質をもつための必要十分条件は

(ロ) “ある m 次元整数要素ベクトル $\boldsymbol{y}$ が存在して $\boldsymbol{c}$ を A と $\boldsymbol{y}$ により $\boldsymbol{c}^{\top}=\boldsymbol{y}^{\top}A$ と表せる”

ことである．

証明　十分性は自明．必要性を示すために，A の各要素の共通分母 d を A に掛けて整数を要素とする行列 $\tilde{A}=d\cdot A$ を作り，それに対応して $\boldsymbol{c}$ の方も d 倍して $\tilde{\boldsymbol{c}}=d\cdot\boldsymbol{c}$ (整数要素ベクトルとは限らない)を作る．このとき性質(イ)は

($\tilde{イ}$) “$\tilde{A}\boldsymbol{x}$ が d の倍数を要素とするベクトルとなるようなすべての $\boldsymbol{x}$ に対して $\tilde{\boldsymbol{c}}^{\top}\cdot\boldsymbol{x}$ は d の倍数である”

と述べ直される．

単模行列 S,T (それぞれ，m 次，n 次)により，$\tilde{A}$ を(6.6.2)の形 $A'=S\tilde{A}T$

にする．

$\boldsymbol{x}=T\boldsymbol{z}$ ($\boldsymbol{z}=T^{-1}\boldsymbol{x}$；$T^{-1}$ も単模行列)，$\boldsymbol{c}'=T^{\mathsf{T}}\tilde{\boldsymbol{c}}$ とおく．$\tilde{A}\boldsymbol{x}$ の要素が d の倍数であることと $S\tilde{A}\boldsymbol{x}$ の要素が d の倍数であることとは等価であることに注意すると，性質(ｲ)は

(ｲ′)　"$e_k z_k$ ($k=1,\cdots,r$)が d の倍数であるような($z_{r+1},\cdots,z_n$ は何でもよい)すべての $\boldsymbol{z}$ に対して $\boldsymbol{c}'^{\mathsf{T}}\cdot\boldsymbol{z}$ が d の倍数である"

と述べ直される．$z_{r+1},\cdots,z_n$ の中のどれか一つ z_{r+k} ($k=1,\cdots,n-r$)だけを勝手な有理数に取り，他のすべての z_k を 0 とおくことにより，(ｲ′)が満たされるためには，$c'_{r+k}=0$ でなければならないことがわかる．ある z_k ($k=1,\cdots,r$)以外がみな 0 であるような $\boldsymbol{z}$ を取ると，"$e_k z_k$ が d の倍数$\Rightarrow c'_k z_k$ が d の倍数"が満たされるためには，"c'_k が e_k の倍数"でなければならない．そこで，$c'_k=y'_k e_k$ (y'_k：整数)を満たす y'_k がある($k=1,\cdots,r$)．さらに，$y'_{r+1}=\cdots=y'_m=0$ とおいてベクトル $\boldsymbol{y}'$ を定めれば，$\boldsymbol{c}'^{\mathsf{T}}=\boldsymbol{y}'^{\mathsf{T}}A'$ である．$\boldsymbol{y}=S^{\mathsf{T}}\boldsymbol{y}'$ (S の単模性から $\boldsymbol{y}$ も整数ベクトル)とおき，右から T^{-1} を掛けてもとに戻せば，$\tilde{\boldsymbol{c}}^{\mathsf{T}}=\boldsymbol{y}^{\mathsf{T}}\tilde{A}$．両辺を d で割って $\boldsymbol{c}^{\mathsf{T}}=\boldsymbol{y}^{\mathsf{T}}A$．■

6.7　有理標準形

K は任意の体であるとする．5.4.4 項(特に注意 5.4.4)で述べたように，n 次正方行列 A は，$A-xI$ の 1 と異なる単因子($K[x]$の上での)が $e_1(x),\cdots,e_h(x)$ であるとき，相似変換によって，コンパニオン行列 C_i の直和：

$$\left.\begin{array}{l} C=C_1\dotplus\cdots\dotplus C_h,\quad C_i：m_i\text{ 次},\\ \varphi_{C_i}(x)=(-1)^{m_i}e_i(x),\quad \varphi_{C_i}(x)\mid\varphi_{C_{i+1}}(x)\quad(i=1,\cdots,h-1),\\ \sum_{i=1}^{h}m_i=n \end{array}\right\}\qquad(6.7.1)$$

という形に変換される．この形は，与えられた A から(有限回の)有理演算により構成され，一意に定まる．これを，A の**有理標準形**という(**Frobenius 標準形**ということもある)．(6.7.1)において，$e_i(x)$の(x に関する)次数は m_i であるから，もちろん，$m_1\leqq m_2\leqq\cdots\leqq m_h$ であるが，次数の勘定から，A の 1 に等しい単因子が $n-h$ 個あることもわかる．各 C_i の特性多項式は(符号を除いて)最小多項式に一致する(→例 5.4.1)．C の最小多項式($=A$ の最小多項式)は $e_h(x)$

であり，(符号を除いて) C_h の特性多項式に等しい．

6.8 Jordan 標準形

$K=\mathbb{C}$ とする．

6.8.1 Jordan 標準形[1]

注意 5.2.7 によれば，n 次正方行列 A は相似変換によって次の形に変換することができる．

$$\left.\begin{aligned}
&J=J_1\dot{+}\cdots\dot{+}J_s,\\
&\qquad J_i : n_i \text{ 次},\quad \sum_{i=1}^{s} n_i=n\,;\\
&J_i=J_{i1}\dot{+}\cdots\dot{+}J_{is_i}\,;\\
&J_{ij}=\begin{bmatrix}\lambda_i & 1 & & & O\\ & \lambda_i & 1 & & \\ & O & \ddots & \ddots & \\ & & & \ddots & 1\\ & & & & \lambda_i\end{bmatrix} n_{ij}=\lambda_i I_{n_{ij}}+N_{ij},\\
&J_{ij} : n_{ij} \text{ 次},\quad n_{i1}\leqq n_{i2}\leqq\cdots\leqq n_{is_i},\quad \sum_{j=1}^{s_i} n_{ij}=n_i.
\end{aligned}\right\} \tag{6.8.1}$$

ここで，$\lambda_1,\cdots,\lambda_s$ は A の相異なる固有値で，A の特性多項式は

$$\varphi_A(x)=(-1)^n\prod_{i=1}^{s}(x-\lambda_i)^{n_i} \tag{6.8.2}$$

である．

6.8.2 項で述べるように，s, n_i, s_i, n_{ij} などの“構造定数”(これを **Jordan 型** (Jordan type) と呼ぶことがある) が，A の有理標準形から決定されるので，A に対して (6.8.1) の形は一意に定まる．この形は，A の **Jordan 標準形** (Jordan canonical form, Jordan normal form) と呼ばれている．相似な行列は同一の Jordan 標準形をもつ．逆も真．

J_{ij} のことを **Jordan ブロック**と呼ぶことがある．

---

1) Jordan という名の人物については注意 4.2.5 の脚注を参照されたい．

6.8.2　有理標準形との関係

(6.8.1)の J_{ij} の特性多項式は $\varphi_{J_{ij}}(x)=(-1)^{n_{ij}}(x-\lambda_i)^{n_{ij}}$ で最小多項式に(符号を除いて)一致する(→例 5.4.2)から，$J-xI$ の単因子(1 も含めて)を $e_1(x),\cdots,e_n(x)$ とすると，各 $e_k(x)$ は

$$e_k(x)=\prod_{i=1}^{s}(x-\lambda_i)^{m_{ki}} \tag{6.8.3}$$

という形をしている．そして，m_{ki} は

$$m_{ki}=n_{i,s_i-(n-k)}, \tag{6.8.4}$$

すなわち，大きい方から $n-k+1$ 番目の n_{ij} の値に等しい($k\leqq n-s_i$ のときは $m_{ki}=0$ とおく)ことがわかる．これらの単因子は，6.7 節の有理標準形 C に対する $C-xI$ の単因子($\varphi_C(x)=\varphi_A(x)$ の因子でもある)を(6.8.3)の形に因数分解したものと一致しなければならない．(6.8.4)は n_{ij} から m_{ki} を定めるという形に書かれているが，この関係を逆に用いると，m_{ki} から n_{ij} を定めることもできる．

このように，A から一意に定まる $A-xI$ の単因子を $\mathbb{C}$ において因数分解することによって(これはもはや有理演算の範囲では一般に不可能であるが)，Jordan 標準形の形が一意に定まることがわかる．

注意 6.8.1　$m_{ki}\leqq m_{k+1,i}$ であること，$\varphi_A(x)=(-1)^n\prod_{i=1}^{s}(x-\lambda_i)^{n_i}$ としたとき $\sum_{k=1}^{n}m_{ki}=n_i$ であることなども明らかであろう．$\psi_A(x)=e_n(x)=\prod_{i=1}^{s}(x-\lambda_i)$ ならば，すべての m_{ki} は 0 か 1 である．そこでこの場合，A は“線形な単因子”をもつと言ったりする．これは，もちろん，A のすべての固有値の代数的多重度が幾何学的多重度に等しい(すなわち，すべての $n_{ij}=1$ で，$n_i=s_i$ である)場合である．

6.8.3　固有空間と一般化固有ベクトルの構造

行列 A を Jordan 標準形 J に相似変換する行列を S とし，$J=S^{-1}AS$ を

$$AS=SJ \tag{6.8.5}$$

と書き換える．

A の固有値 λ_i (に対応する J のブロック J_i)に対応する S の列ベクトル$\{\boldsymbol{e}_1^{(i)},\cdots,\boldsymbol{e}_{n_i}^{(i)}\}$の張る部分空間が A の不変部分空間であることは(6.8.5)，(6.8.1)から明らかである．この部分空間を A の固有値 λ_i に対応する**固有空間**と呼ぶ．

さらに，J_i を構成するブロック J_{ij} に対応する S の列ベクトル$\{\boldsymbol{e}_1^{(i,j)},\cdots,\boldsymbol{e}_{n_{ij}}^{(i,j)}\}$

の張る部分空間も A に関して不変であるが，これらの列ベクトルは

$$Ae_k^{(i,j)}=\lambda_i e_k^{(i,j)}+e_{k-1}^{(i,j)} \qquad (k=n_{ij},n_{ij}-1,\cdots,2), \tag{6.8.6a}$$

$$Ae_1^{(i,j)}=\lambda_i e_1^{(i,j)} \tag{6.8.6b}$$

を満たすので，これは $e_{n_{ij}}^{(i,j)}$ から生成される巡回部分空間(5.4.1項)でもある．$e_1^{(i,j)}$ は A の固有値 λ_i に対応する一つの固有ベクトルであるが，$e_k^{(i,j)}$ $(k=2,\cdots,n_{ij})$ は**一般化固有ベクトル**である(→注意 5.2.8)．(一般化固有ベクトルは**根ベクトル**(root vector)ということもある．)

注意 6.8.2 一般化固有ベクトル $x=e_k^{(i,j)}$ は $(A-\lambda_i I)^p x\neq 0$ $(p=1,\cdots,k-1)$，$(A-\lambda_i I)^k x=0$ という性質を有しているが，一般に，このような性質をもつベクトル x のことを**高さ**(height) k の一般化固有ベクトルと呼ぶ．固有値 λ_i に対応する高さ k 以下の一般化固有ベクトルの張る空間は A の不変部分空間で，その次元は，“λ_i に対応する Jordan ブロックについての $\min(k,n_{ij})$ の和” に等しい．

6.8.4 線形反復

4.4 節で論じた線形反復を再び取り上げる．A の Jordan 標準形 $J=S^{-1}AS$ において，$J_{ij}{}^p$ $(p=1,2,\cdots)$ を作ると，その (k,l) 要素は，直接計算すればわかるように，

$$(J_{ij}{}^p)_{kl}=\begin{cases} 0 & (k>l \text{ のとき}), \\ \lambda_i{}^p & (k=l \text{ のとき}), \\ \lambda_i{}^{p-(l-k)}\dbinom{p}{l-k} & (k<l \text{ のとき}) \end{cases} \tag{6.8.7}$$

である．$p\to\infty$ の極限を考えると，$|\lambda_i|<1$ ならこれらの値は 0 に収束し，$|\lambda_i|>1$ なら発散する．$|\lambda_i|=1$ のときは，$n_{ij}=1$ なら振動(一定値 1 に止まることも含めて)，$n_{ij}\geqq 2$ なら発散する．

$A^p=SJ^pS^{-1}$ であるから，任意のベクトル $x^{(0)}$ から出発する線形反復

$$x^{(\nu+1)}=Ax^{(\nu)} \tag{6.8.8}$$

が 0 に収束する系列 $x^{(0)},x^{(1)},\cdots\to 0$ を作り出すための必要十分条件は，すべての $|\lambda_i|<1$，すなわち A の**スペクトル半径** $\rho(A)$ が 1 より小さいことである．

例 6.8.1 4.4.3 項で述べた Hermite 行列を係数行列とする線形方程式系の **Gauss-Seidel 反復**の収束性の十分条件(ii)を調べる．そのために(不必要に強

い？)次の補助定理を用意する．

補助定理　n 次の Hermite 行列 A，n 次正則行列 B に対して，Hermite 行列

$$C = B + B^{\mathrm{H}} - A \tag{6.8.9}$$

が正定値であるとする．このとき

$$F = I - B^{-1}A \tag{6.8.10}$$

のスペクトル半径が1より小さい($\rho(F) < 1$)ための必要十分条件は，A が正定値であることである．

証明　十分性：F の固有値の一つを λ，対応する固有ベクトルを $\boldsymbol{x}$ ($\neq \boldsymbol{0}$)とする．A が正定値であれば $A = B - BF$ により $A\boldsymbol{x} = (1-\lambda)B\boldsymbol{x} \neq \boldsymbol{0}$ であるから $\lambda \neq 1$．さらに，$\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x} = (1-\lambda)\boldsymbol{x}^{\mathrm{H}}B\boldsymbol{x}$．複素共役転置を取って $\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x} = (1-\bar{\lambda})\boldsymbol{x}^{\mathrm{H}}B^{\mathrm{H}}\boldsymbol{x}$．ゆえに，

$$\frac{\boldsymbol{x}^{\mathrm{H}}(B+B^{\mathrm{H}})\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}} = \frac{1}{1-\lambda} + \frac{1}{1-\bar{\lambda}}, \quad \frac{\boldsymbol{x}^{\mathrm{H}}C\boldsymbol{x}}{\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x}} = \frac{1}{1-\lambda} + \frac{1}{1-\bar{\lambda}} - 1 = \frac{1-|\lambda|^2}{|1-\lambda|^2}.$$

C が正定値，A が正定値であるから $\boldsymbol{x}^{\mathrm{H}}C\boldsymbol{x} > 0$，$\boldsymbol{x}^{\mathrm{H}}A\boldsymbol{x} > 0$ で，$|\lambda|^2 < 1$．これが F のすべての固有値について成り立つから，$\rho(F) < 1$．

必要性：rank $A < n$ なら $A\boldsymbol{x} = \boldsymbol{0}$ を満たす $\boldsymbol{x}$ ($\neq \boldsymbol{0}$)が存在し，$\boldsymbol{x}$ は $F\boldsymbol{x} = \boldsymbol{x}$ を満たすので $\rho(F) \geqq 1$．そこで $\rho(F) < 1$ であるためにはすべての $\boldsymbol{x}$ ($\neq \boldsymbol{0}$)に対して $A\boldsymbol{x} \neq \boldsymbol{0}$．rank $A = n$ でかつ A が正定値でなかったとすると，符号指数 $n_- > 0$ であるので，$\boldsymbol{x}^{(0)\mathrm{H}}A\boldsymbol{x}^{(0)} < 0$ を満たす $\boldsymbol{x}^{(0)}$ が存在する．$\boldsymbol{x}^{(0)}$ ($\neq \boldsymbol{0}$)から出発する F による反復系列 $\boldsymbol{x}^{(\nu+1)} = F\boldsymbol{x}^{(\nu)}$ に注目する．$\rho(F) < 1$ なら，$\nu \to \infty$ のとき $\boldsymbol{x}^{(\nu)} \to \boldsymbol{0}$ であり，したがって $\boldsymbol{x}^{(1)} \neq \boldsymbol{x}^{(0)}$ でもある．またどの ν に対しても $B(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)}) = A\boldsymbol{x}^{(\nu)}$ であるから，

$$(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)})^{\mathrm{H}}B(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)}) = \boldsymbol{x}^{(\nu)\mathrm{H}}A\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)\mathrm{H}}A\boldsymbol{x}^{(\nu)},$$
$$(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)})^{\mathrm{H}}B^{\mathrm{H}}(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)}) = \boldsymbol{x}^{(\nu)\mathrm{H}}A\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu)\mathrm{H}}A\boldsymbol{x}^{(\nu+1)}.$$

これと，$(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)})^{\mathrm{H}}A(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)}) = \boldsymbol{x}^{(\nu)\mathrm{H}}A\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu)\mathrm{H}}A\boldsymbol{x}^{(\nu+1)} - \boldsymbol{x}^{(\nu+1)\mathrm{H}}A\boldsymbol{x}^{(\nu)} + \boldsymbol{x}^{(\nu+1)\mathrm{H}}A\boldsymbol{x}^{(\nu+1)}$ より，$(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)})^{\mathrm{H}}C(\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)}) = \boldsymbol{x}^{(\nu)\mathrm{H}}A\boldsymbol{x}^{(\nu)} - \boldsymbol{x}^{(\nu+1)\mathrm{H}}A\boldsymbol{x}^{(\nu+1)}$．$C$ の正定値性と $\boldsymbol{x}^{(0)} \neq \boldsymbol{x}^{(1)}$ より $(0 >)\,\boldsymbol{x}^{(0)\mathrm{H}}A\boldsymbol{x}^{(0)} > \boldsymbol{x}^{(1)\mathrm{H}}A\boldsymbol{x}^{(1)}$，$\boldsymbol{x}^{(\nu)\mathrm{H}}A\boldsymbol{x}^{(\nu)} \geqq \boldsymbol{x}^{(\nu+1)\mathrm{H}}A\boldsymbol{x}^{(\nu+1)}$ ($\nu = 1, 2, \cdots$)．これは $\boldsymbol{x}^{(\nu)} \to \boldsymbol{0}$, $\boldsymbol{x}^{(\nu)\mathrm{H}}A\boldsymbol{x}^{(\nu)} \to 0$ と矛盾する．ゆえに A は正定値でなければならない．■

A が正定値 Hermite 行列のとき，$A = L + D + L^{\mathrm{H}}$ (L は下三角，D は対角)と

和に分解すると，Gauss-Seidel 反復の反復行列は

$$F=-(L+D)^{-1}L^{\mathrm{H}} \tag{6.8.11}$$

である(→(4.4.26))．A が正定値なら D は正の対角要素よりなる(→注意3.5.14)．$B=L+D$ とおくと，$I-B^{-1}A=I-(L+D)^{-1}(L+D+L^{\mathrm{H}})=F, C=D$ とおけば上の補助定理の条件を満たす．よって $\rho(F)<1$．すなわち，Gauss-Seidel 反復は収束する． □

注意 6.8.3 A が正規行列であれば $\rho(A)=\|A\|_2$ であることはすでに 5.2.2 項で示した．また，注意 5.1.3 では自然ノルムについて $\rho(A)\leqq\|A\|$ が成り立つことを示した．

$\|A\|$ がとくに自然ノルムとは限らないときにも，$\rho(A)\leqq\|A\|$ が成り立つ．実際，$A'=\rho(A)^{-1}A$ とおけば，$\rho(A')=1$，$\|A'\|=\rho(A)^{-1}\|A\|$．もし $\|A\|<\rho(A)$ であったとすると，$\|A'\|<1$．そして，ノルムの性質(→3.3.2 項，(4′))により，$\|A'^p\|\leqq\|A'\|^p$ であるから，$p\to\infty$ のとき $\|A'^p\|\to0$．ノルムの連続性(→3.3.2 項，(1′)〜(3′)と演習問題 3.3)により，このことは $A'^p\to O_{n,n}$ を意味するが，6.8.4 項の始めで見たように，$\rho(A')=1$ であればこのようなことは起きない．

6.8.5 行列多項式，行列関数

2.2 節(注意 2.2.3)，2.12.1，2 項(例 2.12.1，例 2.12.2 も)において行列多項式，行列級数を扱った．A の Jordan 標準形が J である($A=SJS^{-1}$)とき，一変数 x の多項式あるいはべき級数 $f(x)$ に対して

$$f(A)=S\,f(J)\,S^{-1} \tag{6.8.12}$$

であり，J^p は(6.8.7)に示す通り具体的に書き表せるので，Jordan 標準形がわかれば行列多項式や行列関数が具体的に表せることに——原理的には——なる(べき級数の場合には $\rho(A)$ が級数の収束半径より小さくなければならない)．

他方，A の最小多項式 $\psi_A(x)$ の次数が m であれば，$\psi_A(A)=O$，すなわち $A^p\,(p\geqq m)$ は $I, A, A^2, \cdots, A^{m-1}$ の線形結合で表されるから，行列の多項式あるいはべき級数は，$I, A, A^2, \cdots, A^{m-1}$ を用いて表現できるはずである．この辺の事情を以下で調べる．

行列 A の最小多項式が $\psi_A(x)=\prod_{i=1}^{s}(x-\lambda_i)^{m_i}$ であるとき，$\Sigma_A=\{(\lambda_1, m_1), \cdots, (\lambda_s, m_s)\}$ を A の**スペクトル**(spectrum)と呼ぶことがある(要するに，A の固有値とその各々の幾何学的多重度を合わせたもの；5.1 節で定義したスペクトル

sp A とは異なることに注意せよ）．関数 $f(x)$ としては，簡単のため，$\{\lambda_1,\cdots,\lambda_s\}$ を含む複素平面のある領域（連結な開集合）Ω の上で正則な関数のみを対象とする．$f_1(x)$ と $f_2(x)$ が"Σ_A の上で一致する"ということを

各 $\lambda_i\ (i=1,\cdots,s)$ において，f_1 と f_2 の値と m_i-1 階までの導関数の値が一致すること：

$$f_1(\lambda_i)=f_2(\lambda_i)\,,f_1'(\lambda_i)=f_2'(\lambda_i)\,,\cdots,f_1^{(m_i-1)}(\lambda_i)=f_2^{(m_i-1)}(\lambda_i) \qquad (6.8.13)$$

と定義し，$f_1(\Sigma_A)=f_2(\Sigma_A)$ と記す．定義から明らかなように，$f(\Sigma_A)=g(\Sigma_A)$ であれば $h(x)=f(x)-g(x)$ は各 λ_i において m_i 重の零点をもつ，すなわち $h(x)=\psi_A(x)q(x)$（$q(x)$ は正則関数）と表せる．

Ω の上で正則な任意の関数 $f(x)$ に対して，多項式 $g(x)$ で $f(\Sigma_A)=g(\Sigma_A)$ を満たすものが存在するとき，$g(x)$ を Σ_A の上での $f(x)$ の**補間多項式**（polynomial interpolator）という．$g_1(x)$ と $g_2(x)$ が共に $f(x)$ の補間多項式であれば $g_1(x)-g_2(x)$ が $\psi_A(x)$ で割り切れる（逆も真）から，次数を $m-1$ 次以下（$m=\sum_{i=1}^{s}m_i$ は $\psi_A(x)$ の次数）に限れば補間多項式は一意に定まる．

$\psi_A(A)=O$ であるから，$f(x)$ の補間多項式 $g(x)$ として何を選んでも $g(A)$ としては同じ行列が得られる．この行列を $f(A)$ と定義し，関数 $f(x)$ から定まる**行列関数**と呼ぶ．

注意 6.8.4　補間多項式が"存在"することは，具体的には，**Hermite 補間**の公式により示される．すなわち，一般に $f(x)/\psi_A(x)$ は $x=\lambda_i$ に m_i 次の極をもつので

$$\frac{f(x)}{\psi_A(x)}=\sum_{i=1}^{s}\left(\frac{\alpha_{i0}}{(x-\lambda_i)^{m_i}}+\frac{\alpha_{i1}}{(x-\lambda_i)^{m_i-1}}+\cdots+\frac{\alpha_{i,m_i-1}}{x-\lambda_i}\right)+h(x) \qquad (6.8.14)$$

と"部分分数展開"される．ここで，$h(x)$ は Ω で正則であり，α_{ik} は

$$\begin{aligned}\alpha_{ik}&=\frac{1}{k!}\left[\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)^k\frac{f(x)}{\psi_{Ai}(x)}\right]_{x=\lambda_i} \qquad (k=0,1,\cdots,m_i-1),\\ \psi_{Ai}(x)&=\frac{\psi_A(x)}{(x-\lambda_i)^{m_i}}=\prod_{j\neq i}(x-\lambda_j)^{m_j}\end{aligned} \qquad (6.8.15)$$

で与えられる．これにより

$$\begin{aligned}f(x)&=g(x)+\psi_A(x)h(x)\,,\\ g(x)&=\sum_{i=1}^{s}[\alpha_{i0}+\alpha_{i1}(x-\lambda_i)+\cdots+\alpha_{i,m_i-1}(x-\lambda_i)^{m_i-1}]\psi_{Ai}(x)\end{aligned} \qquad (6.8.16)$$

となるから，この $g(x)$ が（$m-1$ 次の）補間多項式であることがわかる．

後で利用するために，(6.8.15) の α_{ik} は $f(\lambda_i),\cdots,f^{(k)}(\lambda_i)$ に，A のみから定まり

f にはよらない係数を掛けて加えた形に書けるので，(6.8.16)の $g(x)$ も

$$g(x)=\sum_{i=1}^{s}\left[f(\lambda_i)\chi_{Ai0}(x)+f'(\lambda_i)\chi_{Ai1}(x)+\cdots+f^{(m_i-1)}(\lambda_i)\chi_{Ai,m_i-1}(x)\right] \tag{6.8.17}$$

という形(ここで，$\chi_{Aik}(x)$ は A (と i と k)のみによって定まる $m-1$ 次の多項式)に表せることにも注意しておこう．

注意 6.8.5 (6.8.17)の $\chi_{Aik}(x)$ を用いて行列

$$Z_{ik}=\chi_{Aik}(A) \qquad (k=0,1,\cdots,m_i-1) \tag{6.8.18}$$

を定義すると，行列関数を“陽な形”(といっても計算は大変であるが)

$$f(A)=\sum_{i=1}^{s}\left[f(\lambda_i)Z_{i0}+f'(\lambda_i)Z_{i1}+\cdots+f^{(m_i-1)}(\lambda_i)Z_{i,m_i-1}\right] \tag{6.8.19}$$

に表せる(例 6.8.2 も見よ)．

注意 6.8.6 $f(x)=1$ の補間多項式は(もとの定義から) $g(x)=1$ であるはずであるから，このとき(6.8.16)は

$$1=\sum_{i=1}^{s}\gamma_{Ai}(x)\psi_{Ai}(x), \qquad \gamma_{Ai}(x)：m_i-1 \text{ 次多項式} \tag{6.8.20}$$

となる．$\psi_{Ai}(x)$ の定義から $\psi_{Ai}(x)\psi_{Aj}(x)\ (i\neq j)$ が $\psi_A(x)$ で割り切れることもわかる．そこで

$$P_i=\gamma_{Ai}(A)\psi_{Ai}(A) \tag{6.8.21}$$

で P_i を定義すると，

$$I=P_1+\cdots+P_s, \qquad P_iP_j=O \qquad (i\neq j) \tag{6.8.22}$$

が得られる．注意 3.7.3 を思い出すと，このことは $P_1,\cdots,P_s$ が射影を表す行列で，これにより全空間 V_n が s 個の部分空間の直和に分けられることがわかる．A を Jordan 標準形 $J=S^{-1}AS$ (→(6.8.1))に表現する基底に関して考えると，$i\neq j$ のとき $\psi_{Ai}(x)$ は $(x-\lambda_j)^{m_j}$ を因子として含むから $\psi_{Ai}(J_j)=O$．すなわち，任意のベクトル $\boldsymbol{x}$ に対して $P_i\boldsymbol{x}$ は固有値 λ_i に対応する固有空間の中にある．したがって，P_i は各固有空間への射影を表す．また，$I=S^{-1}P_1S+\cdots+S^{-1}P_sS$ において $S^{-1}P_iS$ は対角ブロック J_i に対応する次数 n_i の単位行列部分をもつ(その他の部分は O)．$M_i=J_i-\lambda_iI_{n_i}$ とおくと，$M_i^{m_i}=O_{n_i,n_i}$．そこで，もとの基底に戻れば

$$AP_i=\lambda_iP_i+\tilde{M}_i, \quad A=\sum_{i=1}^{s}(\lambda_iP_i+\tilde{M}_i), \quad \tilde{M}_i^{m_i}=O \tag{6.8.23}$$

が得られる($\tilde{M}_i$ は A の次数に達するまで M_i に零行列を加えたものに基底変換を施したもの)．

関数列 $f_1(x), f_2(x), \cdots$ のスペクトル Σ_A 上での“収束”についても，(6.8.13)と同じ精神で，各 λ_i ごとに $\lim_{k\to\infty} f_k^{(p)}(\lambda_i)$ が $p=0,1,\cdots,m_i-1$ に対して存在してそ

れが関数 $f(x)$ の $f^{(p)}(\lambda_i)$ に等しいときに，“f_k は Σ_A 上で f に収束する”ということにする．このとき f_k の補間多項式 g_k も f の補間多項式に収束する．行列関数の収束についても同様．この定義は 2.12.1 項の行列級数の話を特別な場合として含む．

例 6.8.2　Jordan ブロック J_{ij} (→(6.8.1)) の最小多項式は，$\psi_{J_{ij}}(x)=(x-\lambda_i)^{n_{ij}}$ であるから，$f(x)$ の補間多項式は，定義により，

$$g(x)=f(\lambda_i)+\frac{f'(\lambda_i)}{1!}(x-\lambda_i)+\cdots+\frac{f^{(n_{ij}-1)}(\lambda_i)}{(n_{ij}-1)!}(x-\lambda_i)^{n_{ij}-1} \tag{6.8.24}$$

である．$N_{ij}=J_{ij}-\lambda_i I_{n_{ij}}=\begin{bmatrix} 0 & 1 & & & O \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & O & & \ddots & 1 \\ & & & & 0 \end{bmatrix}$ のべき $N_{ij}{}^2, N_{ij}{}^3, \cdots$ は斜めに並ぶ 1 が一つずつ右上にずれるから

$$f(J_{ij})=\begin{bmatrix} f(\lambda_i) & \dfrac{f'(\lambda_i)}{1!} & \cdots & \dfrac{f^{(n_{ij}-1)}(\lambda_i)}{(n_{ij}-1)!} \\ & f(\lambda_i) & \ddots & \vdots \\ & & \ddots & \dfrac{f'(\lambda_i)}{1!} \\ O & & & f(\lambda_i) \end{bmatrix} \tag{6.8.25}$$

という形をしている．これと (6.8.19) を比べると，J_{ij} に対しては

$$Z_{ik}=N_{ij}{}^k \qquad (k=0,1,\cdots,n_{ij}-1) \tag{6.8.26}$$

であることもわかる．注意 6.8.5，注意 6.8.6 と組み合わせると，

$$Z_{i0}=P_i \tag{6.8.27}$$

であることもわかる．

A の Jordan 標準形 $J=S^{-1}AS$ が出来ていれば，J の各ブロックごとに (6.8.25) を作り，それの直和として $f(J)$ を作り，最後に $Sf(J)S^{-1}$ を作れば——原理的には——$f(A)$ が作れることになる．　□

例 6.8.3　2.12.1 項の例に挙げた**指数関数**が任意の正方行列に対して定義できる．A の **Neumann 級数**は，もとになる級数 $(1-x)^{-1}=1+x+x^2+\cdots$ の収束半

径が 1 であるから，A の固有値がすべて $\mathbb{C}$ の単位円の内部にあるとき($\rho(A)<1$ のとき)，そしてそのときに限り，定義される．A の**対数関数**は，もとになる級数 $\log(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\cdots$ の収束半径が 1 であるから，同様にして $\rho(A)<1$ のとき定義される． □

6.8.6 微分方程式，差分方程式

定数係数の 1 階常微分方程式系の初期値問題

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{x}(t)=A\boldsymbol{x}(t), \qquad \boldsymbol{x}(0)=\boldsymbol{x}_0 \tag{6.8.28}$$

を考える．(2.12.26)，(2.12.27) により

$$\frac{\mathrm{d}}{\mathrm{d}t}\exp(tA)=A\cdot\exp(tA), \qquad \exp(O)=I \tag{6.8.29}$$

であるから，

$$\boldsymbol{x}(t)=\exp(tA)\cdot\boldsymbol{x}_0 \tag{6.8.30}$$

が(6.8.28)の解であることがわかる．$\exp(tA)$ は，A の Jordan 標準形ができていれば，例 6.8.2 に従って作ることができる．(という話の筋もあるが，それは「A の固有値の実部がすべて負：$\mathrm{Re}\,\lambda_i<0$ なら，$t\to\infty$ のとき $\boldsymbol{x}(t)\to\boldsymbol{0}$ となる」というような理論的な議論をするのには役に立つものの，具体的に解の形を知りたいときには，A の特性多項式 $\varphi_A(x)$ の零点 λ_i とその多重度 $n_i\,(i=1,\cdots,s)$ を求め，

$$x_j(t)=\sum_{i=1}^{s}(c_{ij0}+c_{ij1}t+\cdots+c_{ij,n_i-1}t^{n_i-1})\mathrm{e}^{\lambda_i t} \tag{6.8.31}$$

という解の“形”を仮定して，未定係数 c_{ijk} を(6.8.28)に代入して定める(微分方程式の方で同一の i に対する c_{ijk} の関係が，初期条件の方で異なる i に対する c_{ijk} の関係が与えられる)という古来の方法を使うのがよい．このやり方だと境界値問題もほとんど同じに扱える．)

定数係数の n 階の差分方程式(簡単のため連立型でないものを考える)の初期値問題

$$\left.\begin{array}{c} x_{\tau+n}=a_{n-1}x_{\tau+n-1}+\cdots+a_1x_{\tau+1}+a_0x_\tau, \\ \tau=0,1,2,\cdots, \\ x_0,x_1,\cdots,x_{n-1}\ \text{が所与} \end{array}\right\} \tag{6.8.32}$$

は，

$$\boldsymbol{x}(\tau)=[x_\tau, x_{\tau+1}, \cdots, x_{\tau+n-1}]^\top \tag{6.8.33}$$

とおくと，

$$\boldsymbol{x}(\tau+1)^\top=\boldsymbol{x}(\tau)^\top C, \quad C=\begin{bmatrix} 0 & \cdots & \cdots & \cdots & 0 & a_0 \\ 1 & \ddots & & & \vdots & a_1 \\ & \ddots & \ddots & & \vdots & \vdots \\ & & \ddots & \ddots & \vdots & \vdots \\ & O & & \ddots & 0 & a_{n-2} \\ & & & & 1 & a_{n-1} \end{bmatrix}, \tag{6.8.34}$$

$\boldsymbol{x}(0)$：所与

と表せるので，(6.8.34)のコンパニオン行列 C のべきを用いて解が

$$\boldsymbol{x}(\tau)^\top=\boldsymbol{x}(0)^\top C^\tau \tag{6.8.35}$$

と表示できる．C^τ は C の Jordan 標準形から例 6.8.2 のやり方で求められる．(これについても，「$\rho(C)<1$ ならば，$\tau\to\infty$ のとき $\boldsymbol{x}(\tau)\to\boldsymbol{0}$」というような“理論”には有用であるが，具体的な解の形を求めたいときには，C の特性多項式 $\varphi_C(x)$ の零点 λ_i と多重度 n_i $(i=1,\cdots,s)$ を計算して

$$x_\tau=\sum_{i=1}^{s}(c_{i0}+c_{i1}\tau+\cdots+c_{i,n_i-1}\tau^{n_i-1})\lambda_i{}^\tau, \qquad \tau=0,1,2,\cdots \tag{6.8.36}$$

という形の解を仮定して，未定係数 c_{ik} を初期条件に代入して決定するという古来の方法を使うのがよい．このやり方は，境界値問題などにも柔軟に対応できる．)

6.8.7　レゾルベントと摂動

複素平面から A の固有値を除いた残りの集合を A の**レゾルベント集合**(resolvent set)と呼び，Res(A)と記すこともある．

x が A の固有値に等しくないところでは(すなわち，$x\in$ Res(A)ならば)，変数 x の有理式を要素とする行列

$$R(A,x)=(A-xI)^{-1} \tag{6.8.37}$$

が定義できる．$R(A,x)$を A の**レゾルベント**(resolvent)と呼ぶ．$R(A,x)$は次の二つの基本的な関係(**レゾルベント等式**(resolvent equations))を満たす．

(i)　$R(A,x)-R(A,y)=(x-y)R(A,x)R(A,y)$, 　　(6.8.38)

(ii)　$R(A,x)-R(B,x)=R(A,x)(B-A)R(B,x)$

$$=R(B,x)(A-B)R(A,x). \tag{6.8.39}$$

証明 (i) $(A-yI)-(A-xI)=(x-y)I$ の両辺に左から $R(A,x)$，右から $R(A,y)$ を掛ければよい．($A,I,R(A,x),R(A,y)$ などが可換なことに注意.)
(ii) $(B-xI)-(A-xI)=B-A$ の左から $R(A,x)$ を掛けると $R(A,x)(B-xI)-I=R(A,x)(B-A)$．さらに右側から $R(B,x)$ を掛ける．第2の等式も同様. ∎

$x\in\mathrm{Res}(A)$ のとき，x の十分小さい近傍の中には A の固有値が存在しないから，その中に y を取れば，(i)により

$$\frac{R(A,y)-R(A,x)}{y-x}=R(A,y)R(A,x).$$

$y\to x$ とすることにより，

$$\frac{\mathrm{d}}{\mathrm{d}x}R(A,x)=R(A,x)R(A,x). \tag{6.8.40}$$

したがって，$R(A,x)$は，レゾルベント集合 $\mathrm{Res}(A)$の上で x の複素関数として微分可能，したがって，x の正則関数である．

注意6.8.4，注意6.8.5において $f(x)=1/(x-y)$ とおくと，$f(A)=R(A,y)$であり，$f^{(k)}(x)=(-1)^k k!/(x-y)^{k+1}$ であるから，

$$R(A,y)=\sum_{i=1}^{s}\sum_{k=0}^{m_i-1}(-1)^k k!\frac{Z_{ik}}{(\lambda_i-y)^{k+1}}. \tag{6.8.41}$$

これに複素関数論の留数定理を適用し，例6.8.2の(6.8.27)と合わせると，A の固有空間への射影 P_i をレゾルベントで表す式

$$P_i=Z_{i0}=-\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_i}R(A,y)\,\mathrm{d}y \tag{6.8.42}$$

が得られる．ここで，Γ_i は，複素平面上で A の固有値 λ_i だけを内部に含み他の固有値はすべて外部にあるような，反時計回りに向きづけられた円(任意の単純閉曲線でもよい)とする．さらに，Z_{ik} $(k=1,\cdots,m_i-1)$ についても

$$Z_{ik}=-\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_i}\frac{(\lambda_i-y)^k}{k!}R(A,y)\,\mathrm{d}y \tag{6.8.43}$$

という表式が得られる．

A の要素がごくわずかずつ変動したときに，固有値，固有空間がどのように変化するかを調べる．十分小さい ε (>0) と一般の n 次正方行列 F を用いて，

A の**摂動**(perturbation)を

$$A(\varepsilon)=A+\varepsilon F \tag{6.8.44}$$

と表しておく(もっと一般の形も想定できるが，ここでは簡単のためこの形の摂動に限る)．$\varepsilon\to 0$ のとき特性多項式は(x の各べきの係数が収束するという意味で)

$$\varphi_{A(\varepsilon)}(x)\to\varphi_A(x)\,. \tag{6.8.45}$$

多項式の零点は係数の連続関数であるから，$A(\varepsilon)$の固有値を，$\varepsilon\to 0$ のとき A のどの固有値 λ_i ($i=1,\cdots,s$；多重度は n_i)に収束するかに従って分類して，

$$\begin{aligned}&\lambda_{11}(\varepsilon),\cdots,\lambda_{1p_1}(\varepsilon) \qquad [\to\lambda_1],\\ &\cdots\cdots\\ &\lambda_{s1}(\varepsilon),\cdots,\lambda_{sp_s}(\varepsilon) \qquad [\to\lambda_s]\end{aligned} \tag{6.8.46}$$

とする(ε が十分小さければ，異なる類に属する $A(\varepsilon)$の固有値は互いに分離していることに注意)．ここで $\lambda_{ij}(\varepsilon)$の多重度を h_{ij} とすれば，

$$h_{i1}+\cdots+h_{ip_i}=n_i \tag{6.8.47}$$

である．注意6.8.6により，$\lambda_{ij}(\varepsilon)$ に対応する $A(\varepsilon)$ の固有空間への射影を $P_{ij}(\varepsilon)$とし，

$$P_i(\varepsilon)=P_{i1}(\varepsilon)+\cdots+P_{ip_i}(\varepsilon) \tag{6.8.48}$$

とおけば

$$I=P_1(\varepsilon)+\cdots+P_s(\varepsilon)\,. \tag{6.8.49}$$

ここで，$A(\varepsilon)$のレゾルベントを考えると，y と A の固有値が δ 以上離れていれば，δ に比べて十分小さいすべての ε に対して $R(A(\varepsilon),y)$が存在し，

$$A(\varepsilon)-yI=A-yI+\varepsilon F=(I+\varepsilon FR(A,y))(A-yI)$$

であるから

$$R(A(\varepsilon),y)=R(A,y)(I+\varepsilon FR(A,y))^{-1}. \tag{6.8.50}$$

ε を十分小さく取れば，$(I+\varepsilon FR(A,y))^{-1}$ の Neumann 級数は($\lambda_1,\cdots,\lambda_s$ を含まない任意の有界閉集合 Ω の中の y に対して一様に)収束し

$$\begin{aligned}R(A(\varepsilon),y)&=R(A,y)(I+\varepsilon FR(A,y))^{-1}\\ &=R(A,y)+\sum_{k=1}^{\infty}\varepsilon^k R^{(k)}(A,y)\,,\end{aligned}$$

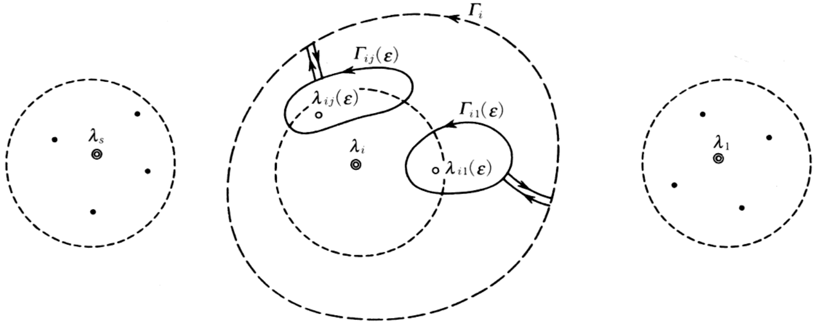


図 6.8.1 $\sum_{j=1}^{p_i}\int_{\Gamma_{ij}(\varepsilon)}=\int_{\Gamma_i}$ の説明

$$R^{(k)}(A,y)=(-1)^k R(A,y)\cdot\underbrace{F\cdot R(A,y)\cdot F\cdot R(A,y)\cdot\cdots\cdot F\cdot R(A,y)}_{k\text{ 組}} \tag{6.8.51}$$

と表せる．各 $\lambda_{ij}(\varepsilon)$ のみを含む反時計回りの閉曲線を $\Gamma_{ij}(\varepsilon)$ とし，$\lambda_{i1}(\varepsilon),\cdots,\lambda_{ip_i}(\varepsilon)$ と λ_i とだけを含む反時計回りの閉曲線を Γ_i とする．（ε が十分小さければ，$\Gamma_{ij}(\varepsilon)$ も Γ_i も Ω の中に選ぶことができる．Γ_i は ε に関係しないように選ぶことができる．）すると，(6.8.42)，(6.8.48) により

$$\begin{aligned}P_i(\varepsilon)&=\sum_{j=1}^{p_i}P_{ij}(\varepsilon)=-\frac{1}{2\pi\mathrm{i}}\left(\sum_{j=1}^{p_i}\int_{\Gamma_{ij}(\varepsilon)}R(A(\varepsilon),y)\,\mathrm{d}y\right)\\&=-\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_i}R(A(\varepsilon),y)\,\mathrm{d}y.\end{aligned} \tag{6.8.52}$$

（図 6.8.1 を見よ；$\lambda_{ij}(\varepsilon)\neq\lambda_i$ なら $R(A(\varepsilon),y)$ は $y=\lambda_i$ において正則であることに注意．）これに (6.8.51) を代入すると

$$\begin{aligned}P_i(\varepsilon)&=P_i+\sum_{k=1}^{\infty}\varepsilon^k P_i^{(k)},\\P_i^{(k)}&=-\frac{1}{2\pi\mathrm{i}}\int_{\Gamma_i}R^{(k)}(A,y)\,\mathrm{d}y\end{aligned} \tag{6.8.53}$$

という，$P_i(\varepsilon)$ の摂動を ε のべき級数で表す式が得られる（$P_i^{(k)}$ は ε に無関係であることに注意）．このことは，

$\varepsilon\to 0$ のとき，$A(\varepsilon)$ の固有値は A の固有値に収束し，$A(\varepsilon)$ の固有空間のうち A の同一の固有値に収束する固有値に対応するものの直和は

A の対応する固有空間に収束する

ということを意味している("部分空間の収束"を定義していないので，この言明は不正確であるが，ここでは，部分空間への射影の表現行列の収束という意味に理解してほしい).

6.9 特異値標準形

$K=\mathbb{C}$ であるとし，V を n 次元計量空間，U を m 次元計量空間，$m\times n$ 行列 A を写像 $A: V\to U$ の表現行列とする．V と U には，それぞれ，正規直交基底が選ばれていて，それに関して内積が，$\boldsymbol{x}_1, \boldsymbol{y}_1\in V$ に対しては $\boldsymbol{x}_1{}^{\mathrm{H}}\cdot\boldsymbol{y}_1$ で，$\boldsymbol{x}_2, \boldsymbol{y}_2\in U$ に対しては $\boldsymbol{x}_2{}^{\mathrm{H}}\cdot\boldsymbol{y}_2$ で表されるとする．V，U の正規直交基底を適当に選んで，A をなるべく簡単な形に表現する問題を扱う．すなわち，n 次ユニタリ行列 Q_1 と m 次ユニタリ行列 Q_2 により A を $Q_2{}^{\mathrm{H}}AQ_1$ と変換することを考える．

結論を先に述べると，任意の A に対して適当な Q_1, Q_2 を選ぶと

$$\Sigma=Q_2{}^{\mathrm{H}}AQ_1=\overbrace{\left[\begin{array}{cccc|c}\sigma_1 & & & O & \\ & \sigma_2 & & & O \\ O & & \ddots & & \\ & & & \sigma_r & \\ \hline & & O & & O\end{array}\right]}^{n}\Bigg\}m, \quad r=\operatorname{rank} A, \tag{6.9.1}$$

$$\sigma_1\geqq\sigma_2\geqq\cdots\geqq\sigma_r>0$$

とすることができる．$\sigma_1,\cdots,\sigma_r$ を A の**特異値**(singular values)，$A=Q_2\Sigma Q_1{}^{\mathrm{H}}$ を A の**特異値分解**(singular-value decomposition；略して SVD)という．Σ は A の**特異値標準形**とでも呼ぶべきものである．

(6.9.1)のような変形ができたとすると

$$\begin{aligned}\Sigma^{\mathrm{H}}\Sigma&=Q_1{}^{\mathrm{H}}A^{\mathrm{H}}AQ_1=\operatorname{diag}[\underbrace{\sigma_1{}^2,\cdots,\sigma_r{}^2,0,\cdots,0}_{n}],\\ \Sigma\,\Sigma^{\mathrm{H}}&=Q_2{}^{\mathrm{H}}AA^{\mathrm{H}}Q_2=\operatorname{diag}[\underbrace{\sigma_1{}^2,\cdots,\sigma_r{}^2,0,\cdots,0}_{m}]\end{aligned} \tag{6.9.2}$$

であるから，$\sigma_1{}^2,\cdots,\sigma_r{}^2$ は非負定値 Hermite 行列 $A^{\mathrm{H}}A$ および AA^{H} の非零固有値である(したがって，一意に定まる)．(rank AA^{H}=rank $A^{\mathrm{H}}A$=rank $A=r$ にも注意.)これを手掛りに，特異値分解の存在と構成法を示すことができる．

n 次非負定値 Hermite 行列 $A^{\mathrm{H}}A$ を対角化するユニタリ行列の一つを Q_1 とする：

$$Q_1{}^{\mathrm{H}}A^{\mathrm{H}}AQ_1=\mathrm{diag}\,[\lambda_1,\cdots,\lambda_r,0,\cdots,0]. \tag{6.9.3}$$

一般性を失うことなく，$\lambda_1 \geqq \cdots \geqq \lambda_r > 0$ としてよい．$Q_1=[\boldsymbol{q}_{\cdot 1}^{(1)},\cdots,\boldsymbol{q}_{\cdot n}^{(1)}]$ とおけば $\{\boldsymbol{q}_{\cdot j}^{(1)} \mid j=1,\cdots,n\}$ は正規直交系をなす．また(6.9.3)により

$$(A\boldsymbol{q}_{\cdot i}^{(1)})^{\mathrm{H}}(A\boldsymbol{q}_{\cdot j}^{(1)})=\boldsymbol{q}_{\cdot i}^{(1)\mathrm{H}}A^{\mathrm{H}}A\boldsymbol{q}_{\cdot j}^{(1)}=\begin{cases}0 & (i \neq j),\\ \lambda_i & (i=j=1,\cdots,r),\\ 0 & (i=j=r+1,\cdots,n)\end{cases} \tag{6.9.4}$$

である．$\sigma_i=\sqrt{\lambda_i}$ $(i=1,\cdots,r)$ とおいて，r 本の m 次元ベクトル $\boldsymbol{q}_{\cdot j}^{(2)}$ $(j=1,\cdots,r)$ を

$$\boldsymbol{q}_{\cdot j}^{(2)}=\frac{1}{\sigma_j}A\boldsymbol{q}_{\cdot j}^{(1)} \qquad (j=1,\cdots,r) \tag{6.9.5}$$

で定義すると，(6.9.3)により，$i,j=1,\cdots,r$ に対しては

$$\boldsymbol{q}_{\cdot i}^{(2)\mathrm{H}}\cdot\boldsymbol{q}_{\cdot j}^{(2)}=\frac{1}{\sigma_i\sigma_j}\boldsymbol{q}_{\cdot i}^{(1)\mathrm{H}}A^{\mathrm{H}}A\boldsymbol{q}_{\cdot j}^{(1)}=\begin{cases}0 & (i \neq j),\\ 1 & (i=j)\end{cases} \tag{6.9.6}$$

であるから $\{\boldsymbol{q}_{\cdot i}^{(2)} \mid i=1,\cdots,r\}$ は U における正規直交系をなす．$r<m$ ならば，適当な $\boldsymbol{q}_{\cdot i}^{(2)}$ $(i=r+1,\cdots,m)$ をこれに付加することによって U の正規直交基底 $\{\boldsymbol{q}_{\cdot i}^{(2)} \mid i=1,\cdots,m\}$ を作ることができる．$Q_2=[\boldsymbol{q}_{\cdot 1}^{(2)},\cdots,\boldsymbol{q}_{\cdot m}^{(2)}]$ とおけば(6.9.1)の分解を得る．

注意 6.9.1 特異値分解は，AA^{H} あるいは $A^{\mathrm{H}}A$ のユニタリ相似(＝合同)変換による対角化と実質的に同内容であるから，その一意性は，対角化をするためのユニタリ行列の一意性による．Q_1 の各列の偏角と，$A^{\mathrm{H}}A$ の等しい固有値群に対応する固有空間の中でのユニタリ変換だけの自由度を Q_1 が有し，Q_1 を定めれば Q_2 の最初の r 列は(6.9.5)により決定されるが，残りの $m-r$ 列は Q_2 がユニタリになるという条件を満たす限り自由に選べる．

注意 6.9.2 特異値分解(6.9.1)は，$A=Q_2\Sigma Q_1{}^{\mathrm{H}}$ と書いてみると，計量空間 V から U への一般の線形写像 A が，(i)まず V の中でのユニタリ変換(鏡映変換，回転)をおこない，(ii)次いで，$\sigma_1,\cdots,\sigma_r,0,\cdots,0$ で定まる各座標軸方向の伸縮をおこない，(iii)最後に U の中でのユニタリ変換をおこなう，という 3 段階の変換に分けて考えられるということを表している．連続体の線形な変形を扱うとき，ユニタリ変換(回転や鏡映)を非本質的な変形とみなすと，どんな変形も，適当な直交座標軸を選べば，軸方向の伸縮として表せるということを意味している．

注意 6.9.3　行列 A の左や右からユニタリ行列を掛けることによって変化しない A の特性量は，A の特異値を用いれば簡明に表すことができる．たとえば，$m \times n$ 型の行列 A に対して

$$\|A\|_2 = \sigma_1, \qquad \|A\|_F = \sqrt{\sigma_1{}^2 + \cdots + \sigma_r{}^2}, \tag{6.9.7}$$

n 次正則行列 A に対して

$$|\det A| = \sigma_1 \cdot \cdots \cdot \sigma_n, \qquad \|A^{-1}\|_2 = 1/\sigma_n, \qquad \mathrm{cond}_2\, A = \sigma_1/\sigma_n, \tag{6.9.8}$$

など．

注意 6.9.4　「任意の n 次正則行列 A は，正定値 Hermite 行列 B とユニタリ行列 Q の積の形 $A=QB$ に表すことができる」というような定理も，特異値分解を用いれば容易に証明できる．実際，(6.9.1)を $A=Q_2 \Sigma Q_1{}^{\mathrm{H}} = Q_2 Q_1{}^{\mathrm{H}} Q_1 \Sigma Q_1{}^{\mathrm{H}}$ と書き直せば $B=Q_1 \Sigma Q_1{}^{\mathrm{H}}$ は正定値 Hermite 行列($r=n$ であるから)，$Q_2 Q_1{}^{\mathrm{H}}$ はユニタリである．

7 一 般 逆 行 列

線形変換 $A: V\to U$ は，その表現行列が正則である(すなわち正方で行列式が 0 でない)とき，そしてそのときに限り，逆変換 $A^{-1}: U\to V$ を有する．それ以外の場合，すなわち，$\dim U\neq\dim V$(表現行列が矩形)であったり，$\dim U=\dim V$(表現行列が正方)であっても表現行列が特異であったりする場合にも，逆変換あるいは逆行列の概念を拡張してみると，応用上役に立つことがある．

一般逆変換(一般逆行列)の理論を展開するには，射影(→3.7節)の概念が本質的な役割を果たす．

特にいちいち注意はしないが，本章の議論はすべて，有限回の有理演算で具体的に辿れるものである．

7.1 一般逆変換と一般逆行列

K は一般の体であるとする．線形写像 $A: V_n\to V_m$ に対して，A の像空間 $U^1\equiv\mathrm{Im}\,A$ は V_m の(次元が $r=\mathrm{rank}\,A$ の)部分空間であり，A の零空間 $W^0\equiv\mathrm{Ker}\,A$ は V_n の(次元が $n-r$ の)部分空間である(→3.2節，3.3節)．

V_m において U^1 の**一つの**補空間 U^0 を選んで固定し，V_n において W^0 の**一つの**補空間 W^1 を選んで固定する．

V_m の基底として U^1 の基底と U^0 の基底の合併を選び，V_n の基底として W^1 の基底と W^0 の基底の合併を選べば，そのような基底に関する A の表現行列は

$$\begin{array}{c|c|c|} & \overset{\longleftarrow r \longrightarrow}{} & \overset{\longleftarrow n-r \longrightarrow}{} \\ \hline r & A_{11} & O \\ \hline m-r & O & O \\ \hline \end{array} \qquad (\det A_{11}\neq 0) \qquad (7.1.1)$$

となる．さらに，W^1 の基底変換と U^1 の基底変換とにより A_{11} の部分を階数標準形に直せば，A の表現行列は

$$\begin{array}{c|c|c|} \multicolumn{1}{c}{} & \multicolumn{1}{c}{\overleftrightarrow{\;r\;}} & \multicolumn{1}{c}{\overleftrightarrow{\;n-r\;}} \\ \cline{2-3} r & \begin{matrix} 1 & & O \\ & \ddots & \\ O & & 1 \end{matrix} & O \\ \cline{2-3} m-r & O & O \\ \cline{2-3} \end{array} = m\begin{array}{|c|c|} \multicolumn{2}{c}{\overleftrightarrow{\quad n \quad}} \\ \hline I_r & O \\ \hline O & O \\ \hline \end{array} \tag{7.1.2}$$

となる．はじめに与えられた V_n, V_m の基底から再出発すれば，A は，

$$SAT = \begin{array}{|c|c|} \hline I_r & O \\ \hline O & O \\ \hline \end{array} \qquad \left(\begin{array}{l} S : m \text{ 次正則} \\ T : n \text{ 次正則} \end{array}\right) \tag{7.1.3}$$

という形に変形できることになる．

いま，V_m から V_n への一つの線形写像 $A^- : V_m \to V_n$ が，A の逆変換としての特徴のうち重要なものの一つ

$$\boldsymbol{y} = A\boldsymbol{x} \quad \text{ならば} \quad \boldsymbol{y} = A(A^-\boldsymbol{y}) \tag{7.1.4}$$

を有しているとする．このことは($n \times m$ 型行列)A^- が

$$AA^-A = A \tag{7.1.5}$$

を満たすことと等価である．A を(7.1.3)の形に表しておくと(7.1.5)を満たす A^- が存在して，その一般形は

$$A^- = T\begin{array}{|c|c|} \hline I_r & D_2 \\ \hline D_3 & D_4 \\ \hline \end{array}S \qquad \left(\begin{array}{l} D_2, D_3, D_4 \text{ は任意で，型はそれぞれ} \\ \quad D_2 : r \times (m-r),\ \ D_3 : (n-r) \times r, \\ \quad D_4 : (n-r) \times (m-r) \end{array}\right) \tag{7.1.6}$$

であることがわかる．

注意 7.1.1　(7.1.5)で特徴づけられる A^- にはどのくらいの本質的な自由度があるかは次のようにして具体的に示される．(7.1.6)において

$$T \text{の代りに}\quad T\begin{array}{|c|c|} \hline I_r & O \\ \hline -D_3 & I_{n-r} \\ \hline \end{array}, \quad S \text{の代りに} \begin{array}{|c|c|} \hline I_r & -D_2 \\ \hline O & I_{m-r} \\ \hline \end{array}S \tag{7.1.7}$$

とおけば，(7.1.6)は

$$A^- = T \begin{array}{|c|c|} \hline I_r & O \\ \hline O & D_4' \\ \hline \end{array} S \qquad (D_4' = D_4 - D_3 D_2) \tag{7.1.6$'$}$$

となる．一方，(7.1.3)の形は変わらない．(7.1.7)は補空間 W^1, U^0 の選び方の変更に対応するので，結局，A^- のもつ自由度は，(i)　V_m における $U^1=\mathrm{Im}\,A$ の補空間 U^0 の選び方，(ii)　V_n における $W^0=\mathrm{Ker}\,A$ の補空間 W^1 の選び方，(iii) D_4'：$U^0 \to W^0$ の選び方，であることがわかる．

このように，A に対して，注意7.1.1の自由度を残して定まる A^- を，A の**一般逆行列**(generalized inverse)と呼ぶ．(7.1.5)が A^- の“定義”であると思ってよい．

(7.1.5)を A^- が満たせば，補空間 U^0, W^1 を

$$U^0 = \mathrm{Ker}\,AA^-, \qquad W^1 = \mathrm{Im}\,A^-A \tag{7.1.8}$$

と選ぶことにより，A^- は(7.1.6′)の形に表される．また，(7.1.5)から

$$(A^-A)(A^-A) = A^-A, \qquad (AA^-)(AA^-) = AA^- \tag{7.1.9}$$

であるから，A^-A は V_n における射影 P_{W^1,W^0} であり，AA^- は V_m における射影 P_{U^1,U^0} であることもわかる(→3.7.1項)．

A が正則であるときは $A^- = A^{-1}$ であることはいうまでもない．

$$\mathrm{rank}\,A^-A = \mathrm{rank}\,AA^- = \mathrm{rank}\,A \tag{7.1.10}$$

であることも明らか．

7.2 最小ノルム形一般逆行列

$K=\mathbb{C}$ とする．A：$V_n \to V_m$．V_n は計量空間であり，計量が $\boldsymbol{x}, \boldsymbol{y} \in V_n$ に対して $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}$ で表されるような(正規直交)基底だけを考えることにする．(V_m には計量が与えられていなくてよい．したがって，V_m の基底は任意でよい．)

rank $A=r$ のとき，与えられた $\boldsymbol{y} \in \mathrm{Im}\,A$ ($\subseteqq V_m$)に対して線形方程式系

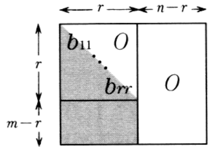
$$A\boldsymbol{x} = \boldsymbol{y} \qquad (\boldsymbol{y} \in \mathrm{Im}\,A) \tag{7.2.1}$$

を満たすベクトル $\boldsymbol{x} \in V_n$ の中で最小の2乗ノルム $\|\boldsymbol{x}\|_2 = \sqrt{\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x}}$ をもつものを定める問題を考える．

A^{H} をQR分解(→2.15.5項)してみればわかる通り，ある n 次ユニタリ行列

Q を用いて A は

$$AQ = $$



$$(b_{ii} \neq 0,\ \ r = \mathrm{rank}\,A) \tag{7.2.2}$$

と変形される．これに行基本変形(→2.15.1項)を施すことによって

$$SAQ = \underbrace{\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}}_{m \times n} \tag{7.2.3}$$

と変形できる．方程式系(7.2.1)を $SAQ(Q^{\mathrm{H}}\boldsymbol{x}) = S\boldsymbol{y} \equiv \tilde{\boldsymbol{y}}$ と書き直せば明らかなように，(7.2.1)を満たす $\boldsymbol{x}$ が存在すれば $\tilde{y}_{r+1} = \cdots = \tilde{y}_m = 0$ で，$\boldsymbol{x}$ は

$$\boldsymbol{x} = Q\tilde{\boldsymbol{x}}, \qquad \tilde{\boldsymbol{x}} = [\tilde{y}_1, \cdots, \tilde{y}_r, \tilde{x}_{r+1}, \cdots, \tilde{x}_n]^{\mathrm{T}} \tag{7.2.4}$$

という形に表せる($\tilde{x}_{r+1}, \cdots, \tilde{x}_n$ は任意)．ところで，$\boldsymbol{x}^{\mathrm{H}}\boldsymbol{x} = \tilde{\boldsymbol{x}}^{\mathrm{H}}\tilde{\boldsymbol{x}}$ であるから，解(7.2.4)の中でノルムが最小のものは $\tilde{x}_{r+1} = \cdots = \tilde{x}_n = 0$ である．このような $\boldsymbol{x}$ は，そこで，

$$A^{\vee} = Q \begin{bmatrix} I_r & D_2 \\ O & D_4 \end{bmatrix} S \qquad \left(\begin{array}{l} D_2: r \times (m-r)\text{型}, \\ D_4: (n-r) \times (m-r)\text{型}, \\ D_2 \text{と} D_4 \text{は任意} \end{array} \right) \tag{7.2.5}$$

とおくことにより，$\boldsymbol{x} = A^{\vee}\boldsymbol{y}$ と表される．

注意 7.2.1　方程式(7.2.1)の解 $\boldsymbol{x}$ に“最小ノルム”の条件を課したものが $\boldsymbol{y}$ の線形式として表せるということは，自明ではない．

$A^{\vee}$ が7.1節で定義した A^{-} の特別な場合であることは明らかである．$A^{\vee}$ を A の**最小ノルム形一般逆行列**(minimum-norm generalized inverse)と呼ぶ．

注意 7.2.2　(7.2.3)，(7.2.4)において Q がユニタリであるということは，7.1節で定義した V_n の部分空間 $W^0 = \mathrm{Ker}\,A$ の補空間 W^1 が“直交補空間”でなければならないことを意味している(注意7.1.1の T の自由度がない)：

$$W^1 = \mathrm{Im}\,A^{\vee}A = (\mathrm{Ker}\,A)^{\perp}. \tag{7.2.6}$$

言い換えれば，V_n における射影 $A^{\vee}A=P_{W^1,W^0}$ は正射影である（→ 3.7.2 項，注意 (3.7.4)）：

$$(A^{\vee}A)^{\mathrm{H}}=A^{\vee}A. \tag{7.2.7}$$

しかし，S の自由度（V_m における $U^1=\mathrm{Im}\,A$ の補空間 U^0 の選択の自由度）は残るので，注意 7.1.1，(7.1.7) と同様にして，U^0 を適当に選べば

$$A^{\vee}=Q\begin{bmatrix} I_r & O \\ O & D_4 \end{bmatrix}S \tag{7.2.5$'$}$$

と表せる（D_4 は $(n-r)\times(m-r)$ 型で任意）．

7.3　最小誤差形一般逆行列

$K=\mathbb{C}$ とする．$A：V_n\to V_m$．V_m は計量空間であり，計量が $\boldsymbol{x},\boldsymbol{y}\in V_m$ に対して $\boldsymbol{x}^{\mathrm{H}}\boldsymbol{y}$ で表されるような（正規直交）基底だけを考えることにする．（V_n には計量が与えられていなくてよい．したがって，V_n の基底は任意でよい．）

rank $A=r$ のとき，与えられた $\boldsymbol{y}\in V_m$ に対して（$\boldsymbol{y}\notin\mathrm{Im}\,A$ であってよい）$\|A\boldsymbol{x}-\boldsymbol{y}\|_2$ が最小になるような $\boldsymbol{x}$ $(\in V_n)$ を求める問題（4.6 節の最小二乗法の問題である）を考える．

議論しかたは 7.2 節とほとんど同じである．m 次ユニタリ行列 Q と n 次正則行列 T を用いて，A を

$$QAT=\begin{bmatrix} I_r & O \\ O & O \end{bmatrix} \tag{7.3.1}$$

と変形する．$\tilde{\boldsymbol{y}}=Q\boldsymbol{y}$，$\tilde{\boldsymbol{x}}=T^{-1}\boldsymbol{x}$ とおけば，$\|A\boldsymbol{x}-\boldsymbol{y}\|_2=\|QAT\tilde{\boldsymbol{x}}-\tilde{\boldsymbol{y}}\|_2$ を最小にするには

$$\tilde{x}_i=\tilde{y}_i\ (i=1,\cdots,r),\quad \tilde{x}_{r+j}\ (j=1,\cdots,n-r)\ \text{は任意} \tag{7.3.2}$$

とすればよいから，

$$A^{\wedge}=T\begin{array}{|c|c|}\hline I_r & O \\ \hline D_3 & D_4 \\ \hline\end{array}\,Q \qquad \begin{pmatrix} D_3: (n-r)\times r\text{ 型}, \\ D_4: (n-r)\times(m-r)\text{ 型}, \\ D_3\text{ と }D_4\text{ は任意} \end{pmatrix} \tag{7.3.3}$$

とおくことにより，求める $\boldsymbol{x}$ は $\boldsymbol{x}=A^{\wedge}\boldsymbol{y}$ と表される．$A^{\wedge}$ は7.1節で定義した A^{-} の特別な場合で，A の**最小誤差形一般逆行列**(minimum-error generalized inverse)と呼ばれる．

注意 7.3.1　注意7.2.1と同様の注意がここでも成り立つ．

注意 7.3.2　7.1節で定義した V_m の部分空間 $U^1=\mathrm{Im}\,A$ の補空間 U^0 は，U^1 の直交補空間でなければならない：

$$U^0=\mathrm{Ker}\,AA^{\wedge}=(\mathrm{Im}\,A)^{\perp}. \tag{7.3.4}$$

すなわち，V_m における射影 $AA^{\wedge}=P_{U^1,U^0}$ は正射影である(→3.7.2項，注意3.7.4)：

$$(AA^{\wedge})^{\mathrm{H}}=AA^{\wedge}. \tag{7.3.5}$$

しかし，T の自由度は残るので，注意7.1.1，(7.1.7)と同様にして，W^1 を適当に選べば

$$A^{\wedge}=T\begin{array}{|c|c|}\hline I_r & O \\ \hline O & D_4 \\ \hline\end{array}\,Q \tag{7.3.3$'$}$$

と表せる(D_4 は$(n-r)\times(m-r)$型で任意)．

7.4　反射形一般逆行列

K は任意の体とする．$A: V_n\to V_m$．7.1節の一般逆行列の定義では，A がその一般逆行列 A^{-} の一般逆行列であるとは限らない．そこで，(7.1.5)の代りに

$$AA^{\times}A=A, \qquad A^{\times}AA^{\times}=A^{\times} \tag{7.4.1}$$

が両方とも成り立つような $A^{\times}: V_m\to V_n$ のことを，A の**反射形一般逆行列**(reflexive generalized inverse)という．

A を(7.1.3)の形に表せば，(7.4.1)の第1式より

$$A^{\times} = T \begin{array}{|c|c|} \hline I_r & D_2 \\ \hline D_3 & D_4 \\ \hline \end{array} S \qquad (7.4.2)$$

の形が得られるが，これが(7.4.1)の第2式を満たすための必要十分条件は，D_2, D_3, D_4 が任意ではなくて，

$$D_4 = D_3 D_2 \qquad (7.4.3)$$

を満たすことである．

注意 7.4.1　注意 7.1.1 と同様の議論により，V_m における $U^1 = \mathrm{Im}\, A$ の補空間 U^0 および V_n における $W^0 = \mathrm{Ker}\, A$ の補空間 W^1 を適当に選べば

$$A^{\times} = T \begin{array}{|c|c|} \hline I_r & O \\ \hline O & O \\ \hline \end{array} S \qquad (7.4.2')$$

と表せる((7.1.6′)，(7.4.3)による)．

注意 7.4.2　(7.4.1)から

$$\mathrm{rank}\, A = \mathrm{rank}\, A^{\times} \qquad (7.4.4)$$

であることがすぐわかる．また，(7.4.2′)，(7.1.3)を比べることによって，$A^{\times}$ が

$$\mathrm{Im}\, A^{\times} A = \mathrm{Im}\, A^{\times}, \qquad \mathrm{Ker}\, A^{\times} = \mathrm{Ker}\, A A^{\times} \qquad (7.4.5)$$

を満たすことがわかる．

逆に，一般逆行列 A^- が(7.4.4)あるいは(7.4.5)(の二つの式の一方)を満たせば A^- は反射形であることもわかる．

7.5　Moore-Penrose 形一般逆行列

$K = \mathbb{C}$ とする．$A : V_n \to V_m$．V_n も V_m も計量空間で，それぞれの基底としては正規直交基底のみを考える．

A の最小ノルム，最小誤差かつ反射形の一般逆行列を，A の **Moore-Penrose**(ムーア-ペンローズ)**形一般逆行列**といい，A^+ と記すことにする．

A の特異値分解(→ 6.9 節；後出の注意 7.5.1 も参照)を

$$\widetilde{A}=Q_2{}^{\mathrm{H}}AQ_1=\left[\begin{array}{cc|c}\sigma_1 & & \\ & \ddots & O \\ & & \sigma_r \\ \hline & O & O\end{array}\right] \quad (\sigma_1,\cdots,\sigma_r>0) \tag{7.5.1}$$

と表してみると，上の定義の条件から $\widetilde{A}^+=Q_1{}^{\mathrm{H}}A^+Q_2$ が

$$\widetilde{A}^+=\left[\begin{array}{cc|c}\sigma_1^{-1} & & \\ & \ddots & O \\ & & \sigma_r^{-1} \\ \hline & O & O\end{array}\right] \tag{7.5.2}$$

という形をしていなければならない．また

$$A^+=Q_1\widetilde{A}^+Q_2{}^{\mathrm{H}} \tag{7.5.3}$$

がそれらの条件を満たすことも明らかである．

このように，A の Moore-Penrose 形一般逆行列 A^+ は常に存在して(ほとんど)一意に定まる．

注意 7.5.1 各 $\boldsymbol{y}\in V_m$ に対して $\|A\boldsymbol{x}-\boldsymbol{y}\|_2$ を最小にする $\boldsymbol{x}$ の全体を X_y とおく．X_y の中で $\|\boldsymbol{x}\|_2$ を最小にする $\boldsymbol{x}$ は一意に定まり，$\boldsymbol{y}$ に $\boldsymbol{x}$ を対応させる写像は線形写像である．実際，A を特異値分解(7.5.1)して，$\tilde{\boldsymbol{x}}=Q_1{}^{\mathrm{H}}\boldsymbol{x}$，$\tilde{\boldsymbol{y}}=Q_2{}^{\mathrm{H}}\boldsymbol{y}$ とおいて問題を述べ直せば，

$$\sum_{i=1}^{n}|\sigma_i\tilde{x}_i-\tilde{y}_i|^2\to\min, \qquad \sum_{i=1}^{n}|\tilde{x}_i|^2\to\min \tag{7.5.4}$$

となるから，

$$\tilde{x}_i=\tilde{y}_i/\sigma_i \quad (i=1,\cdots,r), \qquad \tilde{x}_i=0 \quad (i=r+1,\cdots,n) \tag{7.5.5}$$

が解であること，そして，これが $\boldsymbol{x}=A^+\boldsymbol{y}$ と表せることは明らかである(A^+ をこのようにして定義することもできる)．

注意 7.5.2 上では，A の Moore-Penrose 形一般逆行列 A^+ を，有理演算(と開平)のみでは実行できない操作である A の特異値分解に頼って説明してきた．これは“わかりやすさ”の点で勝れているし，また多くの書物や解説でもそのような説明のしかたが採用されているが，他の形の一般逆行列と同じように，Moore-Penrose 形のものも，有限回の有理演算(と開平)によって作ることができる．

実際，$U^1=\mathrm{Im}\,A$ の“直交補空間” $U^0=(U^1)^{\perp}$(の基底)は有理的に構成できる．$W^0=\mathrm{Ker}\,A$ の直交補空間についても同様である．$V_m=U^1\oplus U^0$，$V_n=W^1\oplus W^0$ という直和分解に対応した正規直交基底に関して，A は

$$Q_2{}^{\mathrm{H}} A Q_1 = \begin{array}{c} \\ r \\ m-r \end{array}\!\!\begin{array}{c} \quad\overleftrightarrow{\;r\;}\;\;\overleftrightarrow{n-r} \\ \left[\begin{array}{c|c} A_{11} & O \\ \hline O & O \end{array}\right] \end{array}, \qquad \det A_{11} \neq 0 \tag{7.5.6}$$

という形をとる．そこで

$$A^{+} = Q_1 \left[\begin{array}{c|c} A_{11}{}^{-1} & O \\ \hline O & O \end{array}\right] Q_2{}^{\mathrm{H}} \tag{7.5.7}$$

とおけばよい(A_{11} を対角化するのに特異値分解の“非有理性”があるだけである)．

注意 7.5.3　いままで述べてきたすべての形の一般逆行列についていえることであるが，それらの“定義”もまた“構成方法”も，実際のデータや数値計算の過程に少しでも雑音(誤差)が入ると，致命的な影響を受けることが多い．このような意味で，一般逆行列の話はまったく“構造不安定”である．

8
非　負　行　列

本章では，$K=\mathbb{R}$ とおき，すべての要素が非負であるような行列を扱う．このような行列は，計量経済学における投入・産出分析において，また，確率過程論における Markov 連鎖の推移確率行列として現れる．さらに，電気回路などのネットワーク理論や数値計算などにおいても，原因に対して結果が(あるいはその逆が)ある種の“単調性”を有している場合によく現れる．

$K=\mathbb{R}$ の上の行列 A(ベクトルも特別な場合である)の要素がすべて非負のとき，A を**非負行列**(nonnegative matrix)という．すべての要素が正のときには**正行列**(positive matrix)という．行列 A,B が同じ型で，$A-B$ が非負行列のとき $A \geqq B$，$A-B$ が正行列のとき $A>B$ と記す(この記法は本章だけで通用するものである)．特別な場合として，非負ベクトル，正ベクトルなどについても同様である．

$K=\mathbb{R}$ の上の行列を扱うわけではあるが，議論の途中で必要に応じて複素数体 $\mathbb{C}$ にまで K を拡大することもある．

8.1　非負行列，M 行列

8.1.1　非負行列，M 行列

n 次正方非負行列 A と正数 μ に対して，次の 7 条件(i)～(vii)が互いに等価であるということが知られている．

(i)　任意の n 次元非負ベクトル $\boldsymbol{b}$ に対して，$(\mu I_n - A)\boldsymbol{x}=\boldsymbol{b}$ を満たす非負ベクトル $\boldsymbol{x}$ が存在する．

(ii)　$(\mu I_n - A)\boldsymbol{x}>\boldsymbol{0}$ を満たす正ベクトル $\boldsymbol{x}$ が存在する．

(iii)　$p\to\infty$ のとき $(\mu^{-1}A)^p \to O_{n,n}$．

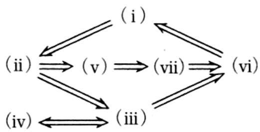
(iv) $\rho(A) < \mu$.

(v) $\mu I_n - A$ のすべての首座小行列式(→ 2.7.1 項)の値が正である．

(vi) $\mu I_n - A$ が正則で $(\mu I_n - A)^{-1} \geqq O_{n,n}$.

(vii) $\mu I_n - A$ は $\mu I_n - A = LU$ (L：対角要素が 1 で非対角要素が非正の下三角行列，U：対角要素が正で非対角要素が非正の上三角行列)と LU 分解される(→ 2.15.3 項)．

以下で上の諸条件の等価性を示す．証明は，



の順でおこなう．

(i)⇒(ii)：$\boldsymbol{b} = [1, 1, \cdots, 1]^\top$ に対して $(\mu I_n - A)\hat{\boldsymbol{x}} = \boldsymbol{b}$ を満たす $\hat{\boldsymbol{x}} \geqq \boldsymbol{0}$ が存在する．$\boldsymbol{e} = [1, \cdots, 1]^\top$ に対して $\boldsymbol{y} = (\mu I_n - A)\boldsymbol{e}$ とおく．$\boldsymbol{y} \geqq \boldsymbol{0}$ のときには $c = 1$，そうでないときには $c = \dfrac{1}{2}\left[\max_i(-y_i)\right]^{-1}$ とおく．すると $\boldsymbol{x} = \hat{\boldsymbol{x}} + c\boldsymbol{e} > \boldsymbol{0}$, $(\mu I_n - A)\boldsymbol{x} = \boldsymbol{b} + c\boldsymbol{y} > \boldsymbol{0}$ である．

(ii)⇒(iii)：$(\mu I_n - A)\hat{\boldsymbol{x}} > \boldsymbol{0}$ を満たす正ベクトル $\hat{\boldsymbol{x}}$ を選び，$\hat{\boldsymbol{y}} = (I_n - \mu^{-1}A)\hat{\boldsymbol{x}}\,(> \boldsymbol{0})$ とおく．$\hat{\boldsymbol{y}} + (\mu^{-1}A)\hat{\boldsymbol{y}} + (\mu^{-1}A)^2\hat{\boldsymbol{y}} + \cdots + (\mu^{-1}A)^p\hat{\boldsymbol{y}} = \hat{\boldsymbol{x}} - (\mu^{-1}A)^{p+1}\hat{\boldsymbol{x}} \leqq \hat{\boldsymbol{x}}$ であるが，この式の左辺は p に関して単調非減少であるから，$(\mu^{-1}A)^p\hat{\boldsymbol{y}} \to \boldsymbol{0}$ $(p \to \infty)$ でなければならない．ところで，$\hat{\boldsymbol{y}} > \boldsymbol{0}$ であるから，$(\mu^{-1}A)^p \to O_{n,n}$ $(p \to \infty)$．

(iii)⟺(iv)：6.8.4 項で示したように，$(\mu^{-1}A)^p \to O_{n,n}\,(p \to \infty)$ の必要十分条件は $\rho(\mu^{-1}A) < 1$，すなわち $\rho(A) < \mu$ である．

(iii)⇒(vi)：$(I_n - \mu^{-1}A)(I_n + \mu^{-1}A + (\mu^{-1}A)^2 + \cdots + (\mu^{-1}A)^p) = I_n - (\mu^{-1}A)^{p+1}$ において，(iii)により，十分大きい p に対して $I_n - (\mu^{-1}A)^{p+1}$ の固有値は 1 に近いから 0 にはならない．ゆえに，そのような p に対して，$I_n - (\mu^{-1}A)^{p+1}$ は正則．したがって，左辺の因子 $I_n - \mu^{-1}A$ も正則，そして $\mu I_n - A$ も正則．さらに，$O_{n,n} \leqq I_n + \mu^{-1}A + (\mu^{-1}A)^2 + \cdots + (\mu^{-1}A)^p = (I_n - \mu^{-1}A)^{-1}(I_n - (\mu^{-1}A)^{p+1})$ において $I_n - (\mu^{-1}A)^{p+1} \to I_n$ であるから，$(I_n - \mu^{-1}A)^{-1} \geqq O_{n,n}$，すなわち，$(\mu I_n - A)^{-1} \geqq O_{n,n}$.

(vi)⇒(i)：$\mu I_n - A$ が正則なら，$(\mu I_n - A)\boldsymbol{x} = \boldsymbol{b}$ は $\boldsymbol{x} = (\mu I_n - A)^{-1}\boldsymbol{b}$ と表せるが，$\boldsymbol{b} \geqq \boldsymbol{0}$, $(\mu I_n - A)^{-1} \geqq O_{n,n}$ なら $\boldsymbol{x} \geqq \boldsymbol{0}$ である．

(ii)⇒(v)：$A_k = A^{1:\cdots:k}_{1:\cdots:k}$ (A の k 次正方部分行列，$1 \leqq k \leqq n$；→ 2.4.1 項)と記す．$(\mu I_n - A)\boldsymbol{x} > \boldsymbol{0}$ を満たす $\boldsymbol{x} = [x_1, \cdots, x_n]^\top > \boldsymbol{0}$ に対して k 次元ベクトル $\boldsymbol{y} = [x_1, \cdots, x_k]^\top > \boldsymbol{0}$ とおくと，$(\mu I_k - A_k)\boldsymbol{y} > \boldsymbol{0}$ (A が非負で $x_{k+1}, \cdots, x_n > 0$ であること

に注意)．(ii)⇒(iv)により，$\rho(A_k)<\mu$．ゆえに，$\det(\mu I_k-A_k)=\prod_{i=1}^{k}(\mu-\lambda_i)$ ($\lambda_1,\cdots,\lambda_k$ は A_k の固有値) において $|\lambda_i|<\mu$．このことから，λ_i が実数なら $\mu-\lambda_i>0$，λ_i が複素数ならそれと複素共役対をなす $\lambda_j=\bar{\lambda}_i$ が存在し，$(\mu-\lambda_i)(\mu-\bar{\lambda}_i)=|\mu-\lambda_i|^2>0$ であることがわかる．したがって，μI_n-A の首座小行列式は $|\mu I_n-A|^{1\cdots k}_{1\cdots k}=\det(\mu I_k-A_k)=\prod_{i=1}^{k}(\mu-\lambda_i)>0$．

(v)⇒(vii)：$B\equiv\mu I_n-A$ の LU 分解 $B=LU$ において，$u_{kk}=|B|^{1\cdots k}_{1\cdots k}/|B|^{1,\cdots,k-1}_{1,\cdots,k-1}$ であるから，(v)により $u_{kk}>0$ ($l_{kk}=1$ は LU 分解の定義)．L,U の非対角要素の符号は，2.15.3 項の LU 分解の算法(A の代りに B とする)において $B^{(k)}$ の非対角要素の符号と一致する．$B^{(1)}=B$ の非対角要素は，A が非負だから，非正である．$B^{(k)}$ の非対角要素が非正なら，L の列 k の行 $k+1,\cdots,n$ の要素および U の行 k の列 $k+1,\cdots,n$ の要素は非正(算法の[2]，上記の $u_{kk}=b_{kk}^{(k)}>0$ にも注意)．$B^{(k+1)}$ の非対角要素は，算法の[3]の $b_{ij}^{(k+1)}:=b_{ij}^{(k)}-l_{ik}u_{kj}$ において，$b_{ij}^{(k)}\leqq 0$，$l_{ik}\leqq 0$，$u_{kj}\leqq 0$ より，$b_{ij}^{(k+1)}\leqq 0$ ($i\neq j$)．

(vii)⇒(vi)：$\mu I_n-A=LU$ において，$u_{kk}>0$ だから $\det(\mu I_n-A)=\prod_{k=1}^{n}u_{kk}>0$，$\mu I_n-A$ は正則．L,U の要素の符号から $L^{-1}\geqq O, U^{-1}\geqq O$．ゆえに，$(\mu I_n-A)^{-1}=U^{-1}L^{-1}\geqq O$．

注意 8.1.1　上記の条件を満たす非負行列 A と正数 μ に対して，A(そして $\mu I-A$)を既約標準形(→ 6.1.1 項)に表しておくと，$(\mu I-A)^{-1}$ の既約対角ブロックは(非負であるだけでなく)正行列となる．とくに，A が既約なら $(\mu I-A)^{-1}>O$．このことは $(\mu I-A)^{-1}=\mu^{-1}(I+\mu^{-1}A+(\mu^{-1}A)^2+(\mu^{-1}A)^3+\cdots)$ と 6.1.1 項で述べたことから明らかであろう．

上記の諸条件を満たす $\mu I-A$ ($A\geqq O,\mu>0$)という形に表すことのできる行列を **M 行列**(M-matrix)と呼ぶ．$B=\mu I-A$ が M 行列であれば，$b_{ij}=-a_{ij}\leqq 0$ ($i\neq j$)で，(v)により $b_{11},\cdots,b_{nn}>0$ (行・列の番号の振り方は本質的でないことに注意)，$\det B>0$．そして(vi)により $B^{-1}\geqq O$ である．逆に，B が正則で，

$$b_{ii}>0\ (i=1,\cdots,n),\quad b_{ij}\leqq 0\ (i\neq j),\quad B^{-1}\geqq O \tag{8.1.1}$$

を満たすとき，

$$\mu=\max_{i=1,\cdots,n} b_{ii},\qquad A=\mu I-B \tag{8.1.2}$$

とおけば，明らかに $\mu>0, A\geqq O$ であって条件(vi)を満たすから，B は M 行列である．したがって，(8.1.1)を M 行列の定義とすることもできる．

例 8.1.1　最も典型的な M 行列は，$m+1$ 個の節点の間を正の線形抵抗のみで

結んだ電気回路(グラフとしては連結グラフであるとする)の**節点コンダクタンス行列**(node conductance matrix)であろう．回路の1点を接地点とし，それを基準にして測った残りの各節点($\alpha=1,\cdots,m$)の電位を v_α と記し，各節点から回路に外から流入する電流を i_α と記す(接地点から流出する電流は $i_1+\cdots+i_m$ となる)．節点コンダクタンス行列は，電位 $v_1,\cdots,v_m$ と電流 $i_1,\cdots,i_m$ を関係づける式

$$i_\alpha=\sum_{\beta=1}^{m} y_{\alpha\beta}v_\beta \qquad (\alpha=1,\cdots,m) \tag{8.1.3}$$

の係数行列 $Y=[y_{\alpha\beta}]$(この場合は対称行列！)である．**Ohm**(オーム)**の法則**(Ohm's law)と **Kirchhoff の電流法則**(Kirchhoff's current law)とにより

$$y_{\alpha\beta}=\begin{cases}\text{節点 }\alpha\text{ と }\beta\text{ を直接結ぶ抵抗のコンダクタンス(抵抗の逆数)の総和に負の符号をつけたもの} & (\alpha\neq\beta\text{ のとき}),\\ \text{節点 }\alpha\text{ に接続されている抵抗のコンダクタンスの総和} & (\alpha=\beta\text{ のとき})\end{cases} \tag{8.1.4}$$

と表されることは容易にわかる．このような電気回路はエネルギーを消費するだけであるから，どのような電圧のかけ方に対しても

$$\sum_{\alpha=1}^{m} i_\alpha v_\alpha=\sum_{\alpha=1}^{m}\sum_{\beta=1}^{m} y_{\alpha\beta}v_\alpha v_\beta>0 \qquad (\boldsymbol{v}\neq\boldsymbol{0}) \tag{8.1.5}$$

でなければならない．このことは Y が正定値であること，したがって上の条件(v)を満たすことを示している(→ 注意 3.5.14)．このようにして，Y が M 行列であることが示される．そうすると，Y は他の6条件をすべて満たす($\mu=\max_\alpha y_{\alpha\alpha}$, $A=\mu I-Y$ とおいて)．たとえば，$Y^{-1}\geqq O$ は，電圧・電流関係の"単調性"

「どこかの節点に外部から電流を注入すれば，すべての節点の電位が上がる(下がることはない)」

を表している． □

8.1.2 Perron-Frobenius の定理

既約な(→ 6.1.1 項) n 次正方非負行列 A に対して次の(1)，(2)，(3)が成り立つ．これは **Perron-Frobenius**(ペロン-フロベニウス)**の定理**と呼ばれている．

(1) $\rho(A)>0$ で，A は正の固有値 $\rho(A)$ とそれに対応する“正の”固有ベクトルをもつ．

(2) $A \geqq B(\geqq O_{n,n})$ なる任意の n 次正方(非負)行列 B に対して，$\rho(A) \geqq \rho(B)$ が成り立つ．$\rho(A)=\rho(B)$ が成り立つのは $A=B$ のときに限る．

(3) $\rho(A)$ は A の単純な固有値である．

証明 (1) 8.1.1 項の条件(ii)，(iv)が等価であることにより，任意の $\alpha>0$ に対して $((\rho(A)+\alpha)I-A)\boldsymbol{x}>\boldsymbol{0}$ を満たす正ベクトル $\boldsymbol{x}>\boldsymbol{0}$ が存在する．一般性を失わずに $\|\boldsymbol{x}\|_2=1$ と仮定しておく．$\alpha_k \to 0$ $(k\to\infty)$ なる正数列 $\alpha_1, \alpha_2, \cdots$ に対する $\boldsymbol{x}$ の列 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots$ は，収束する部分列をもつ($n-1$ 次元球面 $S=\{\boldsymbol{x} \mid \|\boldsymbol{x}\|_2=1\}$ はコンパクトであるから)．この部分列自身を $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots$ と書き，$\boldsymbol{x}_k \to \hat{\boldsymbol{x}}$ $(k\to\infty)$ とする．すると，$\rho(A)+\alpha_k \to \rho(A)$，$\hat{\boldsymbol{x}} \geqq \boldsymbol{0}$，$\hat{\boldsymbol{x}} \neq \boldsymbol{0}$ $(\|\hat{\boldsymbol{x}}\|_2=1)$ であるから，$(\rho(A)I-A)\hat{\boldsymbol{x}} \geqq \boldsymbol{0}$．$\hat{\boldsymbol{y}}=(I+A)^{n-1}\hat{\boldsymbol{x}}$ とおくと，6.1.1 項により $(I+A)^{n-1}$ の全要素が正であるから，$\hat{\boldsymbol{y}}>\boldsymbol{0}$．一方，$(I+A)^{n-1}(\rho(A)I-A)\hat{\boldsymbol{x}}=(\rho(A)I-A)(I+A)^{n-1}\hat{\boldsymbol{x}}=(\rho(A)I-A)\hat{\boldsymbol{y}}$．ここで，もし $(\rho(A)I-A)\hat{\boldsymbol{x}}=\boldsymbol{0}$ でなければ，再び 6.1.1 項により $(\rho(A)I-A)\hat{\boldsymbol{y}}>\boldsymbol{0}$ となるので，8.1.1 項の条件(ii)が($\hat{\boldsymbol{y}}$ に関して)成立するのに，$\mu=\rho(A)$ であって(iv)に反する．したがって，$(\rho(A)I-A)\hat{\boldsymbol{x}}=\boldsymbol{0}$ でなければならない．これは $\hat{\boldsymbol{x}}$ が A の固有値 $\rho(A)$ に対応する固有ベクトルであることを示している．さらに，$(1+\rho(A))^{n-1}\hat{\boldsymbol{x}}=(I+A)^{n-1}\hat{\boldsymbol{x}}>\boldsymbol{0}$ より $\hat{\boldsymbol{x}}>\boldsymbol{0}$ となる．

(2) $A^{\top}$ も非負で既約であり A と同じ固有値をもつから，上の(1)により，固有値 $\rho(A)$ に対応する正の固有ベクトル $\hat{\boldsymbol{z}}>\boldsymbol{0}$ をもつ．B の任意の固有値を λ，それに対応する固有ベクトルを $\boldsymbol{y}=[y_1, \cdots, y_n]^{\top}$ とする．また，$\hat{\boldsymbol{y}}=[|y_1|, \cdots, |y_n|]^{\top}$ とする．$\lambda\boldsymbol{y}=B\boldsymbol{y}$ の両辺の各要素ごとの絶対値を取ることにより，$|\lambda|\hat{\boldsymbol{y}} \leqq B\hat{\boldsymbol{y}}$．ゆえに，$|\lambda|\hat{\boldsymbol{z}}^{\top}\hat{\boldsymbol{y}} \leqq \hat{\boldsymbol{z}}^{\top}B\hat{\boldsymbol{y}} \leqq \hat{\boldsymbol{z}}^{\top}A\hat{\boldsymbol{y}}=\rho(A)\hat{\boldsymbol{z}}^{\top}\hat{\boldsymbol{y}}$，すなわち，$|\lambda| \leqq \rho(A)$，$\rho(B) \leqq \rho(A)$．いま $|\lambda|=\rho(B)=\rho(A)$ であるとすると，$|\lambda|\hat{\boldsymbol{z}}^{\top}\hat{\boldsymbol{y}}=\hat{\boldsymbol{z}}^{\top}B\hat{\boldsymbol{y}}=\hat{\boldsymbol{z}}^{\top}A\hat{\boldsymbol{y}}=\rho(A)\hat{\boldsymbol{z}}^{\top}\hat{\boldsymbol{y}}$ でなければならないが，$\hat{\boldsymbol{z}}>\boldsymbol{0}$ であるから $|\lambda|\hat{\boldsymbol{y}} \leqq B\hat{\boldsymbol{y}} \leqq A\hat{\boldsymbol{y}}$ も等号で成り立たなければならない：$|\lambda|\hat{\boldsymbol{y}}=B\hat{\boldsymbol{y}}=A\hat{\boldsymbol{y}}$．これから，$(1+|\lambda|)^{n-1}\hat{\boldsymbol{y}}=(I+A)^{n-1}\hat{\boldsymbol{y}}>\boldsymbol{0}$ (→ 6.1.1 項)．ゆえに $\hat{\boldsymbol{y}}>\boldsymbol{0}$．$B\hat{\boldsymbol{y}}=A\hat{\boldsymbol{y}}$，$B \leqq A$，$\hat{\boldsymbol{y}}>\boldsymbol{0}$ より $B=A$ を得る．

(3) $B_i=A^{1,\cdots,i-1,i+1,\cdots,n}_{1,\cdots,i-1,i+1,\cdots,n}$ (A から列 i と行 i を除いた $n-1$ 次正方行列)とおく($i=1, \cdots, n$)．A の特性多項式を $\varphi(x)=\varphi_A(x)$，B_i の特性多項式を $\varphi_i(x)=\varphi_{B_i}(x)$ とすると $\varphi'(x)=\mathrm{d}(\det(A-xI))/\mathrm{d}x=-\sum_{i=1}^{n}\varphi_i(x)$ (→ 2.12.2 項，(2.12.18))．A の列 i と行 i の要素をすべて 0 でおきかえた行列を C_i とすると，$O \leqq C_i \leqq A, C_i \neq A$ (C_i は既約でない！)．(2)により $\rho(C_i)<\rho(A)$．明らかに $\rho(C_i)=\rho(B_i)$．したがって，$x \geqq \rho(A)$ ならば $\varphi_i(x) \neq 0$．十分大きい x に対しては $(-1)^{n-1}\varphi_i(x)>0$ であるから，$(-1)^{n-1}\varphi_i(\rho(A))>0$．ゆえに $(-1)^n\varphi'(\rho(A))=\sum_{i=1}^{n}(-1)^{n-1}\varphi_i(\rho(A))>0$．$\varphi(\rho(A))=0$，$\varphi'(\rho(A)) \neq 0$ であるから，$\rho(A)$ は A の単純な固有値である． ■

注意 8.1.2 既約でない n 次正方非負行列 A に対しては, (1), (2)の代りに

(1′) $\rho(A) \geqq 0$ で, A は固有値 $\rho(A)$ とそれに対応する"非負"の固有ベクトルをもつ,

(2′) $A \geqq B \geqq O_{n,n}$ なる任意の n 次正方(非負)行列 B に対して $\rho(A) \geqq \rho(B)$

が成り立つ. このことを示すには, $\tilde{A}_k \to A$ $(k \to \infty)$, すなわち A に収束する"既約"な非負行列 $\tilde{A}_k$ の列(存在する！)を考えればよい.

注意 8.1.3 非負正方行列 A は, 非負ベクトル $\boldsymbol{x}$ を非負ベクトル $\boldsymbol{y}=A\boldsymbol{x}$ に移す. すなわち A: $V_n \to V_n$ は V_n の"第1象限" $\{\boldsymbol{x} \mid x_1, \cdots, x_n \geqq 0, \boldsymbol{x} \neq \boldsymbol{0}\}$ を第1象限に写像する. この写像をベクトルの方向 $\boldsymbol{x}/\|\boldsymbol{x}\|_2$ の写像とみなせば, 球面 $S=\{\boldsymbol{x} \mid \|\boldsymbol{x}\|_2=1\}$ の第1象限からそれ自身への"連続"写像である. S の第1象限はコンパクトであるから, Brouwer(ブロウエル)の不動点定理により少なくとも1個の不動点が存在する. その方向をもつベクトルを $\boldsymbol{x}$ とすれば $A\boldsymbol{x}=\lambda\boldsymbol{x}$, すなわち A は非負ベクトル $\boldsymbol{x}$ を固有ベクトルとしてもつ, という議論もできる.

注意 8.1.4 (既約とは限らない) n 次正方非負行列 $A=[a_{ij}]$ のスペクトル半径 $\rho(A)$ は, 任意の正ベクトル $\boldsymbol{x}=[x_1, \cdots, x_n]^{\top}$ に対して

$$\min_{i=1,\cdots,n} \frac{\sum_{j=1}^{n} a_{ij} x_j}{x_i} \leqq \rho(A) \leqq \max_{i=1,\cdots,n} \frac{\sum_{j=1}^{n} a_{ij} x_j}{x_i}, \tag{8.1.6a}$$

$$\min_{j=1,\cdots,n} \frac{\sum_{i=1}^{n} a_{ij} x_i}{x_j} \leqq \rho(A) \leqq \max_{j=1,\cdots,n} \frac{\sum_{i=1}^{n} a_{ij} x_i}{x_j} \tag{8.1.6b}$$

を満たす. 実際, 注意 8.1.2 により, $\rho(A)$ に対応する非負固有ベクトル $\boldsymbol{y}=[y_1, \cdots, y_n]^{\top}$ を $y_1+\cdots+y_n=1$ を満たすように規格化し(y_i を $y_i/(y_1+\cdots+y_n)$ でおきかえればよい), $\sum_{j=1}^{n} a_{ij} y_j = \rho(A) y_i$ を $i=1, \cdots, n$ について加え合わせれば,

$$\sum_{j=1}^{n} \left(\sum_{i=1}^{n} a_{ij} \right) y_j = \rho(A) \sum_{i=1}^{n} y_i = \rho(A) \tag{8.1.7}$$

を得る. これは $\sum_{i=1}^{n} a_{ij}$ $(j=1, \cdots, n)$ の"平均"が $\rho(A)$ であることを表しているので

$$\min_{j=1,\cdots,n} \sum_{i=1}^{n} a_{ij} \leqq \rho(A) \leqq \max_{j=1,\cdots,n} \sum_{i=1}^{n} a_{ij}. \tag{8.1.8}$$

$X=\mathrm{diag}[x_1, \cdots, x_n]$ とおくと $\rho(XAX^{-1})=\rho(A)$ であり, XAX^{-1} に対する (8.1.8) は (8.1.6b) に一致する. A の代りに $A^{\top}$ を考えれば, $\rho(A^{\top})=\rho(A)$ であるから, (8.1.6a) が得られる.

さらに, A が既約のときには, $\boldsymbol{x}$ として $\rho(A)$ に対応する正ベクトルの固有ベクトルを選べば, (8.1.6a) がすべて等号で成立する. (A が既約なら $A^{\top}$ も既約であるから, $A^{\top}$ について同様にすると, (8.1.6b) がある正ベクトル $\boldsymbol{x}$ によって等号で

成立する.)

注意 8.1.5　$\rho(A)$ が A のスペクトル半径であるから，A の他の固有値の絶対値はすべて $\rho(A)$ 以下であるが，A が既約な非負正方行列で“対角要素 a_{ii} がすべて正”のときには，$\rho(A)$ 以外の A の固有値の絶対値はすべて $\rho(A)$ より“小さい”ことが示せる．実際，A の任意の固有値を λ，A の固有値 $\rho(A)$ に対応する正の固有ベクトルを $\boldsymbol{x}=[x_1,\cdots,x_n]^\top$ とし，注意 8.1.4 におけると同様 $X=\mathrm{diag}[x_1,\cdots,x_n]$，$B=X^{-1}AX$ とおけば，λ は B の固有値でもあり，

$$\rho(A)=\sum_{j=1}^{n} a_{ij}x_j/x_i=\sum_{j=1}^{n} b_{ij} \qquad (i=1,\cdots,n) \tag{8.1.9}$$

と表せる．ここで，**Gershgorin の定理**(5.2.3 項，“弱い形”で十分)によれば，λ は n 個の Gershgorin の円

$$\left.\begin{aligned} C_i&=\{z\in\mathbb{C}\mid |z-b_{ii}|\leqq r_i\},\\ r_i&=\sum_{j\neq i} b_{ij}=\sum_{j=1}^{n} b_{ij}-b_{ii}=\rho(A)-b_{ii} \end{aligned}\right\} \tag{8.1.10}$$

のどれかに含まれる．C_i は $[-(\rho(A)-2b_{ii}),\rho(A)]$ を直径とする円である．$a_{ii}>0$ $(i=1,\cdots,n)$ なら $b_{ii}>0$ $(i=1,\cdots,n)$ であり，C_i は原点を中心とする半径 $\rho(A)$ の円 C_0 とは $z=\rho(A)$ において接するだけで，C_0 の内部に含まれてしまう．($b_{i_0i_0}=\min_i b_{ii}$ とおけば C_{i_0} は他のすべての C_i を包含する.)したがって，$\lambda=\rho(A)$ (これは A が既約だから単純)であるか $|\lambda|<\rho(A)$ であるかのいずれかである．

注意 8.1.6　実は，上の注意 8.1.5 において，A 自身のすべての対角要素が正でなくても，ある p に対して A^p のすべての対角要素が正であれば，A^p に対して上と同様の議論をすることによって，$\lambda^p=\rho(A)^p$ でなければ，$|\lambda^p|<\rho(A)^p$ すなわち $|\lambda|<\rho(A)$ であることが結論される．

既約行列 A に対して，6.1.2 項の周期標準形(→6.1.2 項)を考えることにより，p が十分大きい σ(A の周期)の倍数であれば，A^p の対角要素はすべて正であるから，$|\lambda|<\rho(A)$ でなければ(すなわち，$|\lambda|=\rho(A)$ であれば)，$\lambda^p=\rho(A)^p$，$\lambda^{p+\sigma}=\rho(A)^{p+\sigma},\cdots$．ゆえに $\lambda^\sigma=\rho(A)^\sigma$，すなわち，$|\lambda|=\rho(A)$ なる固有値が σ 個まではある可能性があるということになる．

このような固有値は，既約行列 A の周期が σ のとき，実際に σ 個存在する．このことは次のようにして示せる．A の固有値 $\rho(A)$ に対応する固有ベクトルを $\boldsymbol{x}$ とし，A の周期標準形における行・列の分割 $I_0,\cdots,I_{\sigma-1}$ に対応して $\boldsymbol{x}$ も

$$\boldsymbol{x}=[\boldsymbol{x}_0^\top,\cdots,\boldsymbol{x}_{\sigma-1}^\top]^\top \tag{8.1.11}$$

と分割する．すると，$A\boldsymbol{x}=\rho(A)\boldsymbol{x}$ の関係は

$$A_{q,q+1}\boldsymbol{x}_{q+1}=\rho(A)\boldsymbol{x}_q \tag{8.1.12}$$

と表される($q=0,\cdots,\sigma-1$；$q=\sigma-1$ のときは $q+1=0$ とする)．いま，1 の原始 σ 乗根を

$$\zeta=\exp\left(\frac{2\pi \mathrm{i}}{\sigma}\right) \tag{8.1.13}$$

として，

$$\boldsymbol{x}^{(k)}=[\boldsymbol{x}_0{}^{\top}, \zeta^k\boldsymbol{x}_1{}^{\top}, \zeta^{2k}\boldsymbol{x}_2{}^{\top}, \cdots, \zeta^{(\sigma-1)k}\boldsymbol{x}_{\sigma-1}{}^{\top}]^{\top} \qquad (k=0,1,\cdots,\sigma-1) \tag{8.1.14}$$

とおくと($\boldsymbol{x}^{(0)}=\boldsymbol{x}$ である)，(8.1.12)より

$$A_{q,q+1}\boldsymbol{x}_{q+1}^{(k)}=\zeta^k\rho(A)\,\boldsymbol{x}_q^{(k)} \qquad (q=0,1,\cdots,\sigma-1) \tag{8.1.15}$$

であるので，$\boldsymbol{x}^{(k)}$ は

$$A\boldsymbol{x}^{(k)}=\zeta^k\rho(A)\,\boldsymbol{x}^{(k)} \qquad (k=0,1,\cdots,\sigma-1) \tag{8.1.16}$$

を満たす．このことは，$\zeta^k\rho(A)\,(k=0,\cdots,\sigma-1)$ が A の固有値で $\boldsymbol{x}^{(k)}$ がそれに対応する固有ベクトルであることを示している．

注意 8.1.7 $|\lambda|=\rho(A)$ なる A の固有値が $\rho(A)$ だけしかない場合には，$B=\rho(A)^{-1}A$ とおけば $\rho(B)=1$ が B の固有値で，それ以外の B の固有値の絶対値はみな1より小さい．そこで，B の Jordan 標準形を考えれば明らかなように，B の固有値1に対応する非負の固有ベクトルを $\boldsymbol{x}$，$B^{\top}$ の固有値1に対応する非負の固有ベクトル(B の左固有ベクトル)を $\boldsymbol{y}$ とし，$\boldsymbol{y}^{\top}\boldsymbol{\cdot}\boldsymbol{x}=1$ となるように $\boldsymbol{x}$ (あるいは $\boldsymbol{y}$, あるいは両方)のスケールを調節しておけば

$$p\to\infty \text{ のとき } B^p\to\boldsymbol{x}\boldsymbol{y}^{\top}, \quad \text{すなわち } A^p\sim\rho(A)^p\boldsymbol{x}\boldsymbol{y}^{\top} \tag{8.1.17}$$

である．

$|\lambda|=\rho(A)$ なる固有値が σ 個：$\rho(A)$，$\zeta\rho(A)$，…，$\zeta^{\sigma-1}\rho(A)$ (ζ は1の原始 σ 乗根)ある場合には，同様にして考えると，系列 I，B，B^2，…自身は収束しないが，系列の第 p 項までの平均が作る系列は収束する：

$$p\to\infty \text{ のとき } \quad \frac{I+B+B^2+\cdots+B^{p-1}}{p}\to\boldsymbol{x}\boldsymbol{y}^{\top}. \tag{8.1.18}$$

8.2 確率行列と Birkhoff の定理

8.2.1 確 率 行 列

n 次正方非負行列 $A=[a_{ij}]$ は，各列の要素の和が1に等しいとき**確率行列**[1]と呼ばれる：

$$a_{ij}\geqq 0 \quad (i,j=1,\cdots,n), \quad \sum_{i=1}^{n}a_{ij}=1 \quad (j=1,\cdots,n). \tag{8.2.1}$$

要素の和が1に等しい n 次元非負ベクトル $\boldsymbol{x}=[x_1,\cdots,x_n]^{\top}$ を**確率ベクトル**(stochastic vector)と呼ぶ：

1) 各行の要素の和が1に等しいものを確率行列と呼ぶ方が長い慣行であるが，話は大差ないので本書では一貫してこの慣行に反した定義を採用している．

$$x_i \geqq 0 \quad (i=1,\cdots,n), \qquad \sum_{i=1}^{n} x_i = 1. \tag{8.2.2}$$

このような A と $\boldsymbol{x}$ の積 $A\boldsymbol{x}$ がまた確率ベクトルであること：

$$\sum_{i=1}^{n}\left(\sum_{j=1}^{n} a_{ij}x_j\right)=\sum_{j=1}^{n}\left(\sum_{i=1}^{n} a_{ij}\right)x_j=\sum_{j=1}^{n} x_j=1, \tag{8.2.3}$$

したがって，任意の自然数 p に対して $A^p\boldsymbol{x}$ が確率ベクトルであることは明らかである．

逆に，任意の確率ベクトル $\boldsymbol{x}$ との積 $A\boldsymbol{x}$ が確率ベクトルとなるような非負行列 A は確率行列である．実際，$\boldsymbol{x}$ としてどれか一つの要素だけが1で他がみな0であるような非負ベクトルを取れば，$A\boldsymbol{x}$ が確率ベクトルであるための条件は(8.2.1)そのものとなる．

確率行列は非負行列の特別な場合であるので，注意8.1.4，(8.1.8)によれば，“確率行列 A は固有値 $\rho(A)=1$ を有し，それが絶対値最大の固有値である”ことがわかる．固有値1に対応する A の左固有ベクトルとして $\boldsymbol{y}^{\mathrm{T}}=[1,\cdots,1]$ を取ることができることは，定義(8.2.1)から明らかである．

注意8.1.5～8.1.7で論じたように，A が既約であれば，A の固有値1は単純で，A の固有値1に対応する固有ベクトル $\boldsymbol{x}$ は正ベクトルである．($\sum_{i=1}^{n} x_i=1$(すなわち $\boldsymbol{y}^{\mathrm{T}}\cdot\boldsymbol{x}=1$)を満たすように規格化しておくことにする．)

A の周期が $\sigma=1$ であれば，1以外の固有値の絶対値は1より小さくて，十分大きなすべての p に対して $A^p > O_{n,n}$ となり，$p\to\infty$ のとき

$$A^p \to [\boldsymbol{x},\cdots,\boldsymbol{x}] \tag{8.2.4}$$

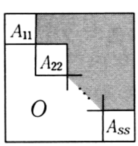
となる．ここで，$[\boldsymbol{x},\cdots,\boldsymbol{x}]$ はすべての列が $\boldsymbol{x}$ に等しい n 次正方行列である．

A が既約で周期 $\sigma(\geqq 2)$ をもつときには，同様にして，$I+A+A^2+\cdots A^{n-1}>O_{n,n}$ で，$p\to\infty$ のとき

$$\frac{I+A+A^2+\cdots+A^{p-1}}{p} \longrightarrow [\boldsymbol{x},\cdots,\boldsymbol{x}] \tag{8.2.5}$$

となる．

A が既約でなくて，行と列の順序を同時に適当に入れ換えたとき

$A=$ 

(8.2.6)

という形に表されたとする(→6.1.1項)．A に付随する有向グラフにおいて，A の対角ブロック A_{kk} に対応する有向グラフの強連結成分が(強連結成分間に定義される半順序関係に関して)極大であれば(すなわち，A_{kk} の“上方”にある A の要素がみな 0 であれば)，A_{kk} 自身が一つの既約な確率行列である．極大でない強連結成分に対応する対角ブロック A_{kk} については，その強連結成分の周期(A_{kk} の周期)が $\sigma=1$ であれば，そのどの列も，十分大きな，すべての p に対して A^p において上方に非零要素が存在する．すなわち，$A^p=[a_{ij}^{(p)}]$ が確率行列であることも考慮に入れると，(8.1.8)より

$$\rho(A_{kk}{}^p)\leqq\max_j\sum_{i\in\text{第}k\text{ブロック}}a_{ij}^{(p)}<1 \tag{8.2.7}$$

であるから，$p\to\infty$ のとき $A_{kk}{}^p\to O$．A_{kk} の周期が $\sigma(\geqq 2)$ のときにも，$(I+A_{kk}+A_{kk}{}^2+\cdots+A_{kk}{}^{p-1})/p$ を考えれば同様である．

例 8.2.1　以上の用語法は，もちろん，**状態**(states) $1,\cdots,n$ をもち，状態 j から i への**推移確率**(transition probability)が a_{ij} で与えられる **Markov**(マルコフ)**連鎖**(Markov chain)のことを念頭においたものである．

ある時点 p において，システムが状態 j にある確率が $x_j(p)$ $(\geqq 0)$ であるとき $\left(\sum_{j=1}^n x_j(p)=1\right)$，次の時点 $p+1$ において状態 i にある確率は

$$x_i(p+1)=\sum_{j=1}^n a_{ij}x_j(p)\ ;\quad \boldsymbol{x}(p+1)=A\boldsymbol{x}(p) \tag{8.2.8}$$

で与えられる．

いままで述べてきた事実を Markov 連鎖の言葉でいい直せば，次の通り．

推移確率行列に付随する有向グラフの極大強連結成分(に対応する状態の集合)を**エルゴード部分**(ergodic　part)，それ以外の強連結成分を**消散部分**(dissipative part)と呼ぶ．どのような初期状態確率 $\boldsymbol{x}(0)$ から始めても，時間 p が経つに従って，$\boldsymbol{x}(p)=A^p\boldsymbol{x}(0)$ であるから，消散部分に属する状態の存在確率は 0 に近づく．エルゴード部分については，各連結成分に対応する状態の類ごとに一つの独立な既約 Markov 連鎖のごとくに振舞い，各類の中での状態確率(あるいはその過去からの平均)は初期状態確率によらない値，すなわち，対応する A の既約成分の固有ベクトルに比例するものに近づく．類の間での確率の分配は初期状態確率に依存する．

8.2.2　二重確率行列と Birkhoff の定理

A も A^{T} も確率行列であるとき，A を**二重確率行列**という．状態 i と j の間の推移が確率的に“対称”であるような場合には，状態間の推移確率が作る行列は対称行列であり，したがって二重確率行列となるが，二重確率行列は対称行列であるとは限らない．

置換行列 $P=[p_{ij}]$ は各行，各列にちょうど1個ずつの1があり，他のすべての要素は0である(→2.3.2項)：

$$p_{ij}=0 \text{ または } 1, \quad \sum_{j=1}^{n} p_{ij}=\sum_{i=1}^{n} p_{ij}=1 \quad (i,j=1,\cdots,n). \tag{8.2.9}$$

したがって，置換行列は特殊な二重確率行列である．ところが，“一般の”二重確率行列は，この特殊な二重確率行列である置換行列の“加重平均”として表せるということ(**Birkhoff**(バーコフ)**の定理**と呼ばれる)が知られている：

「任意の n 次二重確率行列 $A=[a_{ij}]$ に対して，いくつかの置換行列 $P_1=[p_{ij}^{(1)}],\cdots,P_N=[p_{ij}^{(N)}]$ と，正係数 $\alpha_1,\cdots,\alpha_N\ (\alpha_1+\cdots+\alpha_N=1)$ が存在して，

$$A=\alpha_1 P_1+\cdots+\alpha_N P_N \tag{8.2.10}$$

と表される」

ネットワーク・フロー理論(network flow theory)における**最大流問題**(maximum-flow problem)の解法のほとんどは，各辺の容量が整数であるとき整数解を作り出すという性質をもっている．この事実の助けを借りると，Birkhoff の定理は簡単に証明できる．以下にそれを示す．

A の行集合 I に対応する左側頂点集合と列集合 J に対応する右側頂点集合をもち，$a_{ij}\neq 0$ のとき，そしてそのときに限り，頂点 $i(\in I)$ を $j(\in J)$ へ結ぶ辺をもつ**2部グラフ** G (補章A.2節)を考える．G に，さらに入口頂点 s と出口頂点 t を付加し，s を I の各頂点へ結ぶ辺および J の各頂点を t へ結ぶ辺を付加してできる拡大グラフを $\tilde{G}$ とする．

$\tilde{G}$ の各辺に**容量**(capacity)1を与える．$\tilde{G}$ の上の**流れ**(flow)ξ とは，非負実数の組 ξ_{si}，ξ_{ij}，ξ_{jt} $(i,j=1,\cdots,n)$ で，流れの連続条件と容量制約，すなわち

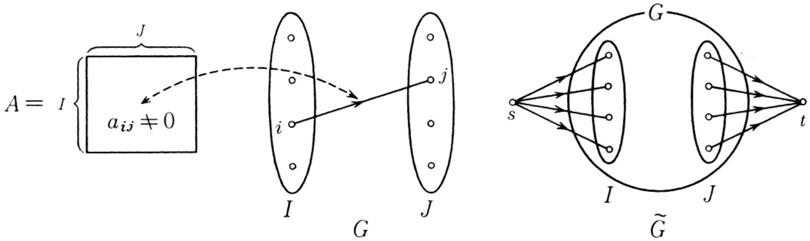


図 8.2.1　二重確率行列 A に付随する 2 部グラフ G とその拡大グラフ $\tilde{G}$

$$\left.\begin{array}{l} \xi_{si}=\sum_{j=1}^{n}\xi_{ij}, \qquad \sum_{i=1}^{n}\xi_{ij}=\xi_{jt}\ ; \\ 1\geqq\xi_{ij}\geqq 0,\quad 1\geqq\xi_{si}\geqq 0,\quad 1\geqq\xi_{jt}\geqq 0\ ; \\ \xi_{ij}=0\,(\text{頂点 } i \text{ を頂点 } j \text{ へ結ぶ辺が存在しないとき}) \end{array}\right\} \tag{8.2.11}$$

を満たすもののことである(直観的意味は明らかであろう)．(8.2.11)から

$$w=\sum_{i=1}^{n}\xi_{si}=\sum_{j=1}^{n}\xi_{jt} \tag{8.2.12}$$

であるが，この w のことを流れ ξ の“総量”という．いまの場合の最大流問題は，(8.2.11)を満たす流れ ξ の中で総量 w が最大であるもの(これを“最大流”という)を求める問題である．最大流問題に関して，次の定理が知られている．(典型的な解法のほとんどはこのような解を与える．)

整数性定理　容量がすべて整数であるとき(いまの場合，容量はすべて 1 であるからもちろん整数である)，$\xi_{si}, \xi_{ij}, \xi_{jt}$ がすべて整数であるような最大流が存在する．

(8.2.11)からは，さらに，

$$w=\sum_{i=1}^{n}\xi_{si}\leqq n, \qquad \max\, w\leqq n \tag{8.2.13}$$

が成り立つことも自明にわかる．

さて，このようにして二重確率行列 A から作ったグラフ $\tilde{G}$ 上の最大流を求めてみよう．

$$\xi_{ij}=a_{ij}, \quad \xi_{si}=1, \quad \xi_{jt}=1, \quad w=n \tag{8.2.14}$$

が“一つの”解であることは自明であろう．そこで，整数性定理により，すべての ξ_{ij} が整数であるような解もある．いまの場合 $1\geqq\xi_{ij}\geqq 0$ であるから，$\xi_{ij}=1$

あるいは0である．$w=n$ であるからすべての $\xi_{si}=\xi_{jt}=1$ である．このことは，

各 i に対して1個そしてただ1個の $\xi_{ij}=1$ が（$a_{ij}>0$ なる(i,j)に対して）あることおよび

各 j に対して1個そしてただ1個の $\xi_{ij}=1$ が（$a_{ij}>0$ なる(i,j)に対して）あることを意味している．これは，このような整数解 $p_{ij}^{(1)}=\xi_{ij}$ から作った行列 $P_1=[p_{ij}^{(1)}]$が置換行列であることを意味している．そこで α_1 を

$$\alpha_1=\min_{p_{ij}^{(1)}=1} a_{ij} \tag{8.2.15}$$

によって定めれば，$A-\alpha_1P_1\geqq O$ であることは明らかである．そして，$A-\alpha_1P_1$ の非零要素の数は，A の非零要素の数より減少している．

$$A^{(1)}=[a_{ij}^{(1)}]=\frac{1}{1-\alpha_1}(A-\alpha_1P_1)\,,\ \text{すなわち}\ A=\alpha_1P_1+(1-\alpha_1)A^{(1)} \tag{8.2.16}$$

とおけば，$A^{(1)}$ は(A より少ない数の非零要素をもつ)二重確率行列である．

以下，同様の議論を繰り返すことによって(8.2.10)の表式が得られる．

9 行　列　束

9.1 行　列　束

$K=\mathbb{C}$ の上の同じ型($m\times n$ 型)の二つの行列 A,B と二つのパラメタ λ,μ に対して，行列の族$\{\mu A-\lambda B\mid\lambda,\mu\in\mathbb{C}\,;\,\lambda=\mu=0$ ではない$\}$のことを A と B から作られる**行列束**(matrix pencil；pencil of matrices)と呼ぶ(以下の話のかなりの部分は K が一般の体でも成り立つ)．行列束を(A,B)のように表す．

$A:V_n\to V_m$，$B:V_n\to V_m$ とみなせば，V_n と V_m のそれぞれの基底の選び方をかえると，表現行列は，m 次正則行列 S，n 次正則行列 T によって

$$A'=SAT,\qquad B'=SBT \tag{9.1.1}$$

という同値変換を受ける．したがって，(9.1.1)で結ばれる(A,B)と(A',B')に対する行列束は**等価**(equivalent)であるとみなせる．

本章では，与えられた二つの行列束(A_1,B_1)，(A_2,B_2)が等価であるかどうかを見分ける方法，および，等価な行列束の **Kronecker 標準形**(Kronecker canonical form)について概観する．

注意 9.1.1　以下で，ときどき，2 変数の同次多項式

$$f(x,y)=a_nx^n+a_{n-1}x^{n-1}y+\cdots+a_1xy^{n-1}+a_0y^n \tag{9.1.2}$$

$(a_i\in K)$を用いる．これは実質的には 1 変数 $\xi=x/y$ の多項式

$$g(\xi)=a_n\xi^n+a_{n-1}\xi^{n-1}+\cdots+a_1\xi+a_0 \tag{9.1.3}$$

と同じ情報しか有しない：

$$f(x,y)=y^ng(x/y)\,. \tag{9.1.4}$$

$a_n=a_{n-1}=\cdots=a_{n-p+1}=0,\ a_{n-p}\neq 0$ のときには，n 次同次多項式 $f(x,y)$ は y^p という因子を有し，$f(x,y)/y^p=f_1(x,y)$ は $n-p$ 次同次多項式となる．

二つの2変数同次多項式 $f_1(x,y)$，$f_2(x,y)$ に対して，f_1 を f_2 で割ったときの商 $q(x,y)$ と剰余 $r(x,y)$ は，x のべきに注目した割算をおこなえば，通常の1変数の多項式の割算と同じようにおこなうことができる：

$$f_1(x,y)=q(x,y)f_2(x,y)+y^p r(x,y). \tag{9.1.5}$$

ただし，$r(x,y)$ は $f_2(x,y)$ より，x のべきに関しては低次であるが，その分 y^p の形の因子が掛かっている．しかし，f_2 が f_1 を割り切るときには $r(x,y)=0$ であるから，このことは問題にならないので，通常の Euclid の互除法は2変数の同次多項式に対して適用可能である．

行列束 (A,B) は

(i) A,B がともに n 次正方行列で $\mu A-\lambda B$ が正則になるような値 μ,λ（μ と λ の比だけが意味がある！）が存在するとき**正則な行列束**(regular pencil of matrices)，

(ii) そうでないとき，すなわち，A,B が $m\times n$ 型で $m\neq n$ のとき，あるいは，$m=n$ でも，すべての μ，λ（$\in\mathbb{C}$）に対して $\mu A-\lambda B$ が特異であるとき，**特異な行列束**(singular pencil of matrices)

と呼ばれる．

(A,B) が正則な行列束のとき，$\det(\mu A-\lambda B)=0$ を満たす (μ,λ) の集合のことを，(A,B) の**スペクトル**(spectrum)という．このとき，$\det(yA-xB)$ は2変数 x,y の同次(n 次)多項式であり，これを行列束 (A,B) の**特性多項式** $\varphi_{(A,B)}(x,y)$ と呼ぶ．

行列束の**階数**(rank)を

$$\operatorname{rank}(A,B)=\max_{\lambda,\mu\in\mathbb{C}}\operatorname{rank}(\mu A-\lambda B) \tag{9.1.6}$$

で定義する．これは，2変数 x,y の多項式行列 $yA-xB$ の階数でもある．

注意 9.1.2 A_1,A_2 が共に n 次正方行列であるとき，正則な行列束 (A_1,I_n) と (A_2,I_n) とが等価であるということは，(9.1.1) より

$$A_2=SA_1T, \qquad I_n=SI_nT \tag{9.1.7}$$

を満たす S,T が存在することであるが，(9.1.7) の第2式は $T=S^{-1}$ を意味するから，A_2 と A_1 が相似であるということに他ならない．(A,I_n) のスペクトルは A のスペクトルに一致する（(A,I_n) のスペクトルが A の固有値の同次座標を与える）．

9.2 一般化固有値問題と正則な行列束

正則な行列束(A,B)において，AあるいはBが，たとえばBが，正定値Hermite行列である場合には，$B=LL^{\mathrm{H}}$(Lは下三角正則行列)と表せる(2.15.4項bおよび演習問題2.6)から

$$A'=L^{-1}A(L^{-1})^{\mathrm{H}}, \qquad I=L^{-1}B(L^{-1})^{\mathrm{H}} \tag{9.2.1}$$

により，(A,B)は$(L^{-1}A(L^{-1})^{\mathrm{H}},I)$と等価である．すなわち，$(A_1,B_1)$，$(A_2,B_2)$に対して，$B_i=L_iL_i^{\mathrm{H}}$，$A_i'=L_i^{-1}A_i(L_i^{-1})^{\mathrm{H}}$とおけば$(A_i,B_i)$は$(A_i',I)$に等価である．したがって，注意9.1.2により，この場合の行列束の等価性や標準形の問題は，第5章，第6章で扱った，単一の行列の相似変換のもとでの等価性や標準形の問題に帰着される(このとき，AがHermite行列ならA'もHermite行列であることに注意されたい)．この場合の問題は伝統的には，**一般化固有値問題**(generalized eigenvalue problem)と呼ばれていた．

正則な行列束(A,B)において，AあるいはBが，たとえばBが，正則行列である(正定値Hermiteとは限らない)場合には，(9.1.1)において$T=I$，$S=B^{-1}$ ($T=B^{-1}$，$S=I$でも同様)とおくことにより，(A,B)は$(B^{-1}A,I)$に等価であることがわかる．したがって，この場合にも，第5章，第6章で扱った単一の行列の場合に問題が帰着される．

注意 9.2.1 (A,B)が正則な行列束のとき，2変数x,yの1次同次式を要素とする行列$yA-xB$のすべてのk次小行列式の最大公約多項式としての**行列式因子**$d_k(x,y)$を考えることができる(→2.15.6項，注意9.1.1)．($d_0(x,y)=1$とする．) $d_k(x,y)|d_{k+1}(x,y)$であり，**単因子**$e_k(x,y)=d_k(x,y)/d_{k-1}(x,y)$が定義される(そして，$e_k(x,y)|e_{k+1}(x,y)$)．$(A,B)$の特性多項式は

$$\varphi_{(A,B)}(x,y)=d_n(x,y)=\prod_{k=1}^{n}e_k(x,y) \tag{9.2.2}$$

であり，$\psi_{(A,B)}(x,y)=e_n(x,y)$を$(A,B)$の**最小多項式**と呼ぶことができる．

これらの定義(の拡張)が意味をもつのは，容易に確かめられるように，それらが等価な行列束に対しては変わらないからである．

上で扱った特殊な正則行列束の場合には，単因子が等しければ等価であることも明らかである．

上で扱った場合以外の正則な行列束(A,B)(A も B も特異)について調べる．

正則性の定義により，$\varphi_{(A,B)}(x,y)=\det(yA-xB)$を 0 にしない変数の値 $x=\lambda_0$, $y=\mu_0$ がある．一般性を失わずに，$\mu_0=1$ であるとする．このとき，変数を

$$\begin{aligned} x&=-\lambda_0 x'+y', \qquad y=-x', \\ x'&=-y, \qquad y'=x-\lambda_0 y \end{aligned} \tag{9.2.3}$$

により x,y から x',y' に変換すると，

$$yA-xB=-y'B-x'(A-\lambda_0 B). \tag{9.2.4}$$

ここで，

$$\begin{aligned} &A_1=-B, \quad B_1=A-\lambda_0 B(=\mu_0 A-\lambda_0 B), \\ &A=-\lambda_0 A_1+B_1, \quad B=-A_1 \end{aligned} \tag{9.2.5}$$

とおくと，行列束(A_1,B_1)の B_1 が正則であるから，上で扱った場合となる．$B_1^{-1}A_1$ を相似変換により Jordan 標準形にしたときの形を $H_0\dot{+}H_1$とする．ここで，H_0 は，$B_1^{-1}A_1$ の固有値 0 に対応する Jordan ブロックを集めた(対角ブロックとして並べた)もので，対角要素の一つ上に斜めに並ぶ要素のいくつかだけが 1 で他がみな 0 の，(上三角)べき零行列である．H_1 は対角要素に非零が並び，その一つ上に斜めに 1 がとびとびに並ぶ以外はすべての要素が 0 の上三角行列(正則)である．(A_1,B_1)をこのように$(H_0\dot{+}H_1,I)$に等価に変換しておいてから，変数 x'，y' を x，y にもどすと，

$$y'(H_0\dot{+}H_1)-x'I=y(I-(\lambda_0 H_0\dot{+}\lambda_0 H_1))-x(-(H_0\dot{+}H_1)). \tag{9.2.6}$$

ここで

$$A_2=I-(\lambda_0 H_0\dot{+}\lambda_0 H_1), \qquad B_2=-(H_0\dot{+}H_1) \tag{9.2.7}$$

とおくと，行列束(A_2,B_2)がはじめの(A,B)と等価である(実際，(A_1,B_1)を$(H_0\dot{+}H_1,I)$に変換するのと同じ行列を左右から掛けることによって，(A,B)が(A_2,B_2)に変換されている)．H_0 が n_0 次，H_1 が n_1 次$(n=n_0+n_1)$であるとする．A_2 と B_2 の左上の n_0 次の対角ブロック $I_{n_0}-\lambda_0 H_0$ と $-H_0$ は ($I_{n_0}-\lambda_0 H_0$ が正則であるから) 左から ($(I_{n_0}-\lambda_0 H_0)^{-1}$ を掛けることによって，それぞれ，I_{n_0} と H_0' (n_0 次の上三角べき零行列) に変換される．残りのブロック $I_{n_1}-\lambda_0 H_1$ と $-H_1$ は $(-H_1)^{-1}$ を掛けることによって，K (上三角行列) と I_{n_1} に変換される．(I_{n_0},H_0') の部分はさらに H_0' の相似変換により(I_{n_0},J_0) (J_0 は対角要素が 0 の Jordan ブロックがいくつか並んだもの)に，(K,I_{n_1}) の部分は同様にして (J_1,I_{n_1}) (J_1 は Jordan ブロックがいくつか並んだもの；対角要素には，一般には，非零も 0 もありうる)に変換される．

このようにして，結局，正則な行列束(A,B)は$(I_{n_0}\dot{+}J_1,J_0\dot{+}I_{n_1})$ (J_0 は対角要素が 0 の Jordan ブロック がいくつか並んだ n_0 次の上三角行列，J_1 は Jordan ブロッ

ク がいくつが並んだ n_1 次の上三角行列，$n_0+n_1=n$）という“標準形”に等価であることがわかる．

注意 9.2.2　上の形が“標準形”の名に値するのは，それが(A, B)に対して一意に定まり，等価な行列束は同じ標準形をもつからである．$\mu A-\lambda B$ の単因子が等価な行列束に対しては同一であること（→注意 9.2.1），$(I_{n_0}\dot{+}J_1, J_0\dot{+}I_{n_1})$に対しては，$(I_{n_0}, J_0)$の部分から生じる単因子が μ^p という形をしていること，(J_1, I_{n_1})の部分からは λ を真に含む単因子が生じることに着目すれば，(A, B)の単因子を求めれば標準形が確定することがわかる．

注意 9.2.3　いうまでもないことであるが，上の標準形において，J_1 の部分で対角要素が 0 である部分を左側に残し，残りのブロックを右側に移す（逆行列の Jordan 標準形を作って）ことも可能である．（A と B，あるいは μ と λ，について話が対称であることに注意！）

9.3　一般の行列束の標準形

行列束(A, B)が正則である場合は既に 9.2 節で扱ったから，本節では特異な行列束について述べる．

行列 A, B は $m\times n$ 型で，

$$m\leqq n,\qquad \mathrm{rank}(A, B)\equiv r<m \tag{9.3.1}$$

であると仮定して一般性を失わない．以下では話の筋道だけを述べ，証明の詳細には立ち入らない．

$yA-xB$ の階数が r であるから

$$(yA-xB)\boldsymbol{x}=\boldsymbol{0} \tag{9.3.2}$$

には $n-r$ 個の線形独立な解 $\boldsymbol{x}$ がある．この場合，$\boldsymbol{x}$（の各要素）は 2 変数 x, y の同次“多項式”であるとする（“有理式”で解をまず表しておいて，“通分”すればよい）．このような解の中で，変数 x についての次数が最も小さいものを $\boldsymbol{x}_1(x, y)$，それに独立で次数が最小なものを $\boldsymbol{x}_2(x, y)$，…というようにして $\boldsymbol{x}_1(x, y)$，$\boldsymbol{x}_2(x, y)$，…，$\boldsymbol{x}_{n-r}(x, y)$が作られる．それらの x に関する次数を $p_1, \cdots, p_{n-r}$ とする．もちろん，$p_1\leqq p_2\leqq\cdots\leqq p_{n-r}$ であるが，はじめのいくつかの p_i は 0 であることもある．そこで $0=p_1=\cdots=p_s<p_{s+1}\leqq\cdots\leqq p_{n-r}$ とする．ここで，$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_{n-r}$ は一意には定まらないが，$p_1, \cdots, p_{n-r}$ は一意に定まる．また，等価な行列束に対しては，この $p_1, \cdots, p_{n-r}$ は等しい．同様にして

$$\boldsymbol{y}^{\mathsf{T}}(yA-xB)=\mathbf{0}^{\mathsf{T}} \tag{9.3.3}$$

を満たす，$\boldsymbol{y}_1(x,y)$，…，$\boldsymbol{y}_{m-r}(x,y)$とその次数の系列 q_1，…，q_{m-r} $(0=q_1=\cdots=q_t<q_{t+1}\leqq\cdots\leqq q_{m-r})$を定める．さらに，(9.3.4)，(9.3.5)により $L_p^{(i)}, M_q^{(i)}$ を定める．

$$L_p^{(1)}=\underbrace{\begin{pmatrix} 0 & 1 & & & O \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ O & & & 0 & 1 \end{pmatrix}}_{p+1}\Bigg\}\,p,\qquad L_p^{(2)}=\underbrace{\begin{pmatrix} 1 & 0 & & & O \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ O & & & 1 & 0 \end{pmatrix}}_{p+1}\Bigg\}\,p; \tag{9.3.4}$$

$$M_q^{(1)}=\underbrace{\begin{pmatrix} 0 & & & O \\ 1 & 0 & & \\ & 1 & \ddots & \\ & & \ddots & 0 \\ O & & & 1 \end{pmatrix}}_{q}\Bigg\}\,q+1,\qquad M_q^{(2)}=\underbrace{\begin{pmatrix} 1 & & & O \\ 0 & 1 & & \\ & 0 & \ddots & \\ & & \ddots & 1 \\ O & & & 0 \end{pmatrix}}_{q}\Bigg\}\,q+1. \tag{9.3.5}$$

$yA-xB$ の(1 および 0 と異なる)単因子から正則な行列束のとき作ったのと同様な“標準形”を作りそれを(A_0, B_0)とする．すると，

$$\hat{A}=\underbrace{\begin{pmatrix} O_{t,s} & & & & & & & O \\ & L^{(1)}_{p_{s+1}} & & & & & & \\ & & \ddots & & & & & \\ & & & L^{(1)}_{p_{n-r}} & & & & \\ & & & & M^{(1)}_{q_{t+1}} & & & \\ & & & & & \ddots & & \\ & & & & & & M^{(1)}_{q_{m-r}} & \\ O & & & & & & & A_0 \end{pmatrix}}_{n}\Bigg\}\,m, \tag{9.3.6 a}$$

$$\hat{B}=\underbrace{\begin{pmatrix} O_{t,s} & & & & & & & O \\ & L^{(2)}_{p_{s+1}} & & & & & & \\ & & \ddots & & & & & \\ & & & L^{(2)}_{p_{n-r}} & & & & \\ & & & & M^{(2)}_{q_{t+1}} & & & \\ & & & & & \ddots & & \\ & & & & & & M^{(2)}_{q_{m-r}} & \\ O & & & & & & & B_0 \end{pmatrix}}_{n}\Bigg\}\,m \tag{9.3.6 b}$$

が，(A,B)の **Kronecker 標準形**と呼ばれるものである．

$(\hat{A},\hat{B})$が(A,B)と等価であって，等しい次数系列 p_1，…，p_{n-r}，q_1，…，q_{m-r}，および等しい単因子をもつことが確認できる．

A
行列式と Pfaffian に対する組合せ論的接近法

A.1 有向グラフ

以下の概念の定義も，各種概念の間に成り立つ関係も，ほとんど自明なものが多いが，グラフ理論の用語は必ずしも統一されていないので，本書の以下の部分での用法を明確にするため基本的な概念の定義を与え，用語の整理をしておく．

抽象的かつ厳密な定義は専門書に譲ることにして，本文 6.1 節におけると同じように，**有向グラフ**(directed graph あるいは oriented graph)(あるいは単に**グラフ**) G とは，**頂点**(vertex，複数形 vertices)(あるいは単に**点**)と二つの点を結ぶ**辺**(edge)(あるいは**有向辺**(arc))とからなる図形(というよりその構造)のことであるとする．

点の(有限)集合を $V=\{v_1, v_2, \cdots\}$ などと記し，辺の(有限)集合を $E=\{e_1, e_2, \cdots\}$ などと記す．辺 e_i が点 v_j を点 v_k に結ぶとき，すなわち e_i が v_j から出て v_k に終わる辺であるとき，v_j を e_i の**始点**(initial vertex)，v_k を e_i の**終点**(terminal vertex)という(図 A.1.1)．始点あるいは終点のことを**端点**(end vertex)ともいう．また，辺はその端点に**接している**(be incident to)ともいう．始点と終点とが一致している辺のことを**ループ**(loop)という(図 A.1.2)．同一の始点と終点をもつ二つ(以上)の辺を**平行辺**(parallel edges)といい，接している辺の存在しない点のことを**孤立点**(isolated vertex)というが，本書では平行辺や孤立点のあるグラフは，特に断らない限り考えないことにする．

どの辺にも一意的にその始点と終点が対応づけられている；すなわち，関数 $\partial^+ : E \to V$，$\partial^- : E \to V$ が定義されている(集合 V と E および関数 ∂^+ と ∂^- が有向グラフ G を定義するというのが，厳密な言い方かもしれない)．

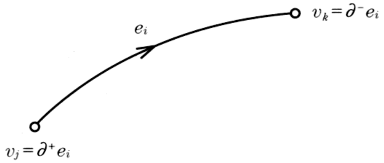


図 A.1.1　辺 e_i の始点 $\partial^+ e_i$ と終点 $\partial^- e_i$



図 A.1.2　ループ e_i

始点関数 ∂^+, 終点関数 ∂^- に対して，それぞれ，逆関数 $\delta^+ : V \to 2^E$, $\delta^- : V \to 2^E$ を考えることができる(2^E は集合 E の部分集合の集合，E のべき集合)．$\delta^+ v$ は点 v を始点としてもつ辺の集合，$\delta^- v$ は点 v を終点としてもつ辺の集合である．$|\delta^+ v| = \mathrm{d}^+ v$, $|\delta^- v| = \mathrm{d}^- v$ と記して，それぞれ点 v の**出次数**(でじすう；outdegree)，**入次数**(いりじすう；indegree)と呼び，$\mathrm{d}v = \mathrm{d}^+ v + \mathrm{d}^- v$ を単に点 v の**次数**(degree)と呼ぶことがある．(v が孤立点なら，$\delta^+ v = \delta^- v = \emptyset$, $\mathrm{d}v = 0$ である．)逆関数の性質から

$$\bigcup_{v \in V} \delta^+ v = \bigcup_{v \in V} \delta^- v = E, \quad \sum_{v \in V} \mathrm{d}^+ v = \sum_{v \in V} \mathrm{d}^- v = |E| \tag{A.1.1}$$

であるが，(A.1.1)の第 2 式から

$$\sum_{v \in V} \mathrm{d}v = 2|E| \tag{A.1.2}$$

を導いて，これを「点の次数の総計は偶数(＝辺の総数の 2 倍)である」と読めば，何となく“定理”の雰囲気がする．(実際(A.1.2)を“グラフ理論の最初の定理”と呼ぶ人がいる[1]．)

点の集合 V, 辺の集合 E, 始終点関数 $\partial^\pm (: E \to V)$ で定められるグラフ $G = (V, E, \partial^+, \partial^-)$ に対して，E の部分集合 $E' (\subseteq E)$ を指定したとき，**辺部分グラフ**(edge subgraph) $G'_{E'}$ とは，点の集合 $V' = V$, 辺の集合 E', $\partial'^+ = \partial^+|_{E'}$, $\partial'^- = \partial^-|_{E'}$ により定義される有向グラフであるとする．普通の言葉でいえば，G から辺の部分集合 $E - E'$ を取り除いたものである．このとき，もし V' の中に孤立点が生じれば，それらは V' から取り除くことにしてもよい．

同様にして，点の部分集合 $V' (\subseteq V)$ を指定して，両端点が共に V' に属すような辺の部分集合を $E' (\subseteq E)$ として，点の集合 V', 辺の集合 E' からなる**点部**

---

1)　Leonhard Euler(レオンハルト・オイラー)がグラフ理論の嚆矢とされる論文の中で最初にあげた命題であるのでこう呼ぶのだという人もいる．また，どのグラフ理論の本にも一番はじめにこの命題が定理としてあげられているからこう呼ぶのだという説もある．

分グラフ(vertex subgraph；induced subgraph ということもある) $G_{V'}''$ を考えることもある．

有向グラフ $G=(V,E,\partial^+,\partial^-)$ 上の**道**(path)とは $P=(v(0),e(1),v(1),e(2),\cdots,v(r-1),e(r),v(r))$ という点と辺とが交互に並んだ系列($v(i)\in V(i=0,1,\cdots,r)$, $e(j)\in E(j=1,\cdots,r)$)で，各 $j=1,\cdots,r$ に対して

(a)　$v(j-1)=\partial^+e(j)$,　$v(j)=\partial^-e(j)$

あるいは

(b)　$v(j-1)=\partial^-e(j)$,　$v(j)=\partial^+e(j)$

であるようなもののことである．ここで，(a)の場合には辺 $e(j)$ は道 P に**順方向に含まれる**といい，(b)の場合には**逆方向に含まれる**という．P に含まれる辺の数(重複も数えた延べ数) r を，道 P の**長さ**(length)という．P に含まれるすべての辺が順方向に含まれるとき，P を**有向道**(directed path)という．有向道の場合には，G に並列辺が存在しないと仮定すれば，有向道 P はその点の系列 $(v(0),\cdots,v(r))$ を与えれば確定する(条件を満たす辺が存在しないこともありうるが，そのような場合を除き)．$v(0)$のことを道 P の始点，$v(r)$のことを道 P の終点といい，それぞれ ∂^+P, ∂^-P と記すこともある．

$\partial^+P=\partial^-P$ であるような道を**閉路**(closed path)，それが有向道であるとき**有向閉路**(directed closed path)という．閉路あるいは有向閉路 P は，その始点＝終点を特別視しないで，点と辺の系列を“円順列”のようにみなすのが普通である．たとえば，点の系列で $P=(v(0),v(1),\cdots,v(r))$ と表される有向閉路については，$v(0)=v(r)$ であるから，$P=(v(1),\cdots,v(r))=(v(2),v(3),\cdots,v(r),v(1))=(v(3),\cdots,v(r),v(1),v(2))=\cdots$ などと表す．特に，辺 e がループであるとき，$v=\partial^+e=\partial^-e$ であるとすれば，$P=(v)$ は長さ 1 の有向閉路とみなせる．

このように閉路あるいは有向閉路については表記に自由度があるので，$P=(v(1),\cdots,v(r))$ と記すときには $v(1)$の V における番号が最小のものを**先頭**(head)に記すことにしておくと便利なことが多い．(もちろん，$m=2,\cdots,r$ のどれかに対して $v(1)=v(m)$ であることもあるから，このような先頭記法は一意に定まらないこともある．)

道あるいは閉路 P に含まれる点がすべて異なるとき(閉路の場合には $v(0)=v(r)$ だけは許す)，P は**初等的**(elementary)であるという．初等的な有向閉路

のことを単に**サイクル**(cycle)と呼ぶ．(ループは長さ1のサイクルでもある．)サイクルに対しては，先頭は一意に定まる(初等性により)．

一般に，有向道 $P=(v(0), v(1), \cdots, v(r-1), v(r))$ が初等的でないときには，$v(k)=v(k')$ ($k<k'$；$k=0$ かつ $k'=r$ ではないとする)であるような k，k' が存在するので，そのような k，k' の中で最小のものを選べば(すなわち，$k=0, 1, 2, \cdots$ と順に調べて行って，はじめて $v(k)=v(k')$ となるものを選べば)，$C=(v(k), v(k+1), \cdots, v(k'-1))$ はサイクルで，P から C を除いたもの，すなわち，$P'=(v(0), \cdots, v(k-1), v(k)=v(k'), v(k'+1), \cdots, v(r))$ もまた $\partial^+P'=\partial^+P$，$\partial^-P'=\partial^-P$ である一つの有向道である．P' の長さは P の長さより短い．そのような操作を繰り返せば，任意の有向道 P は $\partial^+P=\partial^+\tilde{P}, \partial^-P=\partial^-\tilde{P}$ であるような初等的な有向道 $\tilde{P}$ といくつかのサイクルとに分解される．初等的でない有向道を初等的な有向道といくつかのサイクルとに分解するというだけなら，そのとき生じるサイクルは一般には一意に定まらないが，上記の〰〰の条件をつければ当然分解は一意に定まる[1]．(図 A.1.3 を見よ．)

特に P が有向閉路の場合には，P はいくつかのサイクルに分解される．

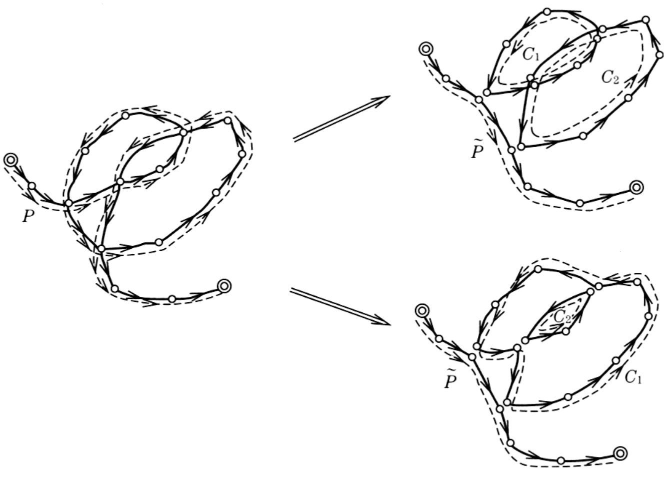


図 A.1.3 初等的でない有向道 P の初等的有向道 $\tilde{P}$ とサイクル C_1, C_2 への分解

1) 不必要な複雑化を避けるため，平行辺はないものとしておく．

道 P は，その中に同一の辺が複数回含まれることが，一般には，許されるが，同一の辺がたかだか1回しか含まれないような場合には，P は**単純**(simple)であるといわれる．初等的な道は単純であることは明らかであろう．

有向閉路が存在しないような(したがってサイクルが存在しないような)有向グラフのことを**無サイクル有向グラフ**(directed acyclic graph)と呼ぶ．近頃はこれを **DAG** と略称することが多いようである．

DAG G の二つの点 $u, v \in V$ の間に $u=\partial^+P$，$v=\partial^-P$ であるような(初等的な)有向道が存在するとき

$$u \rightsquigarrow v \tag{A.1.3}$$

と記すことにする．すべての点 $u \in V$ に対して $u \rightsquigarrow u$ であると約束する(同一の点は長さ0の有向道で結ばれていると解釈する)．すると，V の上にこのようにして定義される関係 $\rightsquigarrow$ は**半順序**(partial order)である．

注意 A.1.1 念のため確認しておくと，以下の3条件1)，2)，3)を満たす V 上の関係 $\rightsquigarrow$ は半順序と呼ばれる．

1) すべての $u(\in V)$ に対して $u \rightsquigarrow u$.
2) $u \rightsquigarrow v$，$v \rightsquigarrow w$ ならば $u \rightsquigarrow w$.
3) $u \rightsquigarrow v$，$v \rightsquigarrow u$ なら $u=v$.

いまの場合は1)，2)，3)が成り立つことは次のようにして確認できる．1)は約束ごと；2)は，$u=\partial^+P_1$，$v=\partial^-P_1$，$v=\partial^+P_2$，$w=\partial^-P_2$ なる有向道 P_1, P_2 が存在すれば，P_1 の列の後に P_2 の列を繋いでできる列が表す有向道 P_3 は $u=\partial^+P_3$ を $w=\partial^-P_3$ に結ぶものである；3)は，2)により，もし $u \neq v$ であれば P_3 は u から始まって v を通り u に戻る有向閉路となるが，考えているグラフが DAG であるからそのような有向閉路は存在しない．

G が DAG ならば，半順序 $\rightsquigarrow$ に関して最上位にあるような点，別の言い方をすれば，それを終点とするような辺が存在しないような点が――一般には複数個――存在する．それらは G の**極大点**(maximal vertices)である．それらの点とそれらを端点(この場合は始点)とする辺を G から取り除いてできる G の辺部分グラフ G_1 を作ると，G_1 もまた DAG である．(サイクルのない G から何かを取り除いたものにサイクルが生じるはずがないから．)このようなことを繰り返せば，G の点の集合 V は，極大点，その次に上位の点，…というように階層化さ

れる．同様にして，**極小点**(minimal vertices)も定義され，それらから順々に点の集合の階層化ができる．しかし，これら二つの階層化は一般には一致しない．

一般に DAG G の点の集合 V を

$$V = V_1 \cup \cdots \cup V_N \qquad (j \neq k \text{ のとき } V_j \cap V_k = \emptyset) \tag{A.1.4}$$

と互いに素な部分集合に分割して，

$$u \in V_j,\ v \in V_k \quad (j \geqq k) \quad \text{ならば} \quad u \rightsquigarrow v \text{ とはならない} \tag{A.1.5}$$

というようにできる．これを，DAG G の点の"関係"$\rightsquigarrow$に即した**階層化**(topological ordering という奇妙な名称も一部で使われている[1])と呼ぶ．階層化は，上にも述べたように一意には定まらないが，それを一つに固定したような DAG のことを**層グラフ**(layered DAG)と呼ぶこともできよう．

DAG に極大点が複数個あるときには，人為的に 1 点 $\hat{u}\,(\notin V)$を付加し，$\hat{u}$を始点とし各極大点を終点とする辺($\delta^+\hat{u}$ と記すことができる)を付け加えれば，最大点を有する DAG を作ることができる．

ほとんど自明に近いことではあるが，最大点 $\hat{u}$ を有する DAG G に対しては以下のような計算が容易にできることに注意しよう．G の各辺 $e\,(\in E)$には重み $w(e)$が付随させられているとする．このとき，$\hat{u}$ から各点 $v\,(\in V)$への有向道 P に沿って辺の重み $w(e)$の積を作り，それらを P について総和したもの，すなわち

$$S(v) = \sum_{\substack{P:\,\partial^+P=\hat{u} \\ \partial^-P=v}} \Bigl(\prod_{\substack{e:\,P\text{ に} \\ \text{含まれる辺}}} w(e) \Bigr) \tag{A.1.6}$$

は，次のようにして計算できる．

[1] 点の集合を $V = V_1 \cup \cdots \cup V_N$ と層化する．

[2] 各点 $v\,(\in V)$に対して作業用の変数 $W(v)$(点 v のポテンシャルなどと呼ぶこともある）を用意する．

[3] すべての v に対して $W(v) := 0$ とおく；
$W(\hat{u}) := 1$ とする．V_1 の各点 v に対しても $W(v) := 1$ とする．

[4] $k = 1, 2, \cdots, N-1$ の順に
V_k の各点 v について

1) たとえば，オペレーションズ・リサーチの日程計画の分野の一つ PERT，CPM などで．

すべての $e(\in \delta^+ v)$ について

$$W(\partial^- e):=W(\partial^- e)+W(v)\cdot w(e). \tag{A.1.7}$$

注意 A.1.2　なお，(A.1.6)，(A.1.7)において，加算($+$やΣ)の代りに最小値を取る演算(min や Min)を，乗算($\cdot$やΠ)の代りに加算を取ると，

$$S(v)=\underset{\substack{P:\partial^+P=\hat{u}\\ \partial^-P=v}}{\mathrm{Min}}\Bigl(\sum_{e:P\text{に含まれる辺}} w(e)\Bigr) \tag{A.1.6$'$}$$

を計算するのには，まずすべての $v(\in V)$ に対して $W(v):=\infty$ とおき，$W(\hat{u}):=0$，すべての $v\in V_1$ に対して $W(v):=0$ として，次に $k=1,2,\cdots,N-1$ の順に，V_k の各点 v について，すべての $e(\in \delta^+ v)$ について，

$$W(\partial^- e):=\min(W(\partial^- e),W(v)+w(e)) \tag{A.1.7$'$}$$

とすればよい．これは，$w(e)$を辺 e の“長さ”とみなしたときの $\hat{u}$ から各点 v までの有向道に沿っての“最短距離”$W(v)$の計算方法を与えているとみなすことができる．(なお，本文第 2 章演習問題 2.1，2.2 も見よ．)

また，上の計算に要する演算の総回数は，加算，乗算がそれぞれ約 $|E|$ 回であることも明らかであろう．(もちろん DAG G の構造を表現するためのデータ構造を適切に設計した上での話ではあるが．)

有向閉路が存在する，したがってサイクルが存在するような有向グラフ G の点の集合 V の上にも，形式的には(A.1.3)の関係 $\rightsquigarrow$ を定義することができる．($u\rightsquigarrow v$ は u から v への有向道が存在することを意味し，すべての $u(\in V)$ に対して $u\rightsquigarrow u$ であると約束する．)これは注意 A.1.1 の条件 1)と 2)は満足するが，3)は満足しない．すなわち，$\rightsquigarrow$ は半順序ではない．そこで，

$$2\text{点 } u,v(\in V)\text{ に対して } u\rightsquigarrow v \text{ かつ } v\rightsquigarrow u \text{ のとき } u\leftrightsquigarrow v \tag{A.1.8}$$

と記すことにすると，

1′)　すべての $u(\in V)$ に対して $u\leftrightsquigarrow u$，

2′)　$u\leftrightsquigarrow v$，$v\leftrightsquigarrow w$ なら $u\leftrightsquigarrow w$，

3′)　$u\leftrightsquigarrow v$ なら $v\leftrightsquigarrow u$

であることは $\leftrightsquigarrow$ の定義から明らかである．これら 3 条件を満たす関係は，集合 V の上の**同値関係**(equivalence relation)と呼ばれ，集合 V はその関係によって**同値類**(equivalence classes) $\tilde{V}_1,\cdots,\tilde{V}_S$ に分割される：

$$V=\tilde{V}_1\cup\cdots\cup\tilde{V}_S \qquad (k\neq l \text{ なら } \tilde{V}_k\cap\tilde{V}_l=\emptyset). \tag{A.1.9}$$

各類 $\tilde{V}_k$ は $u \leftrightsquigarrow v$ であるような u, v が集まって作られたものである．このことを

$$\{\tilde{V}_1, \cdots, \tilde{V}_S\} = V/\leftrightsquigarrow \qquad \text{(A.1.10)}$$

と記したりする．この $\tilde{V}_1, \cdots, \tilde{V}_S$ を有向グラフ G の**強連結成分**(strongly connected components, strong components)と呼ぶ．強連結成分の個数が1であるような有向グラフは**強連結**(strongly connected)であるという．

強連結成分 $\tilde{V}_k, \tilde{V}_l$ に対して，$\tilde{V}_k$ に属する点 u から $\tilde{V}_l$ に属する点 v への有向道が存在する，すなわち $u \rightsquigarrow v$ $(u \in \tilde{V}_k,\ v \in \tilde{V}_l)$のとき $\tilde{V}_k \rightsquigarrow \tilde{V}_l$ と記し，$\tilde{V}_k$ が $\tilde{V}_l$ より大きいとみなすことにする(点の間の関係と同じ記号 $\rightsquigarrow$ を用いるが誤解の恐れはないであろう)．このとき $\tilde{V}_k \rightsquigarrow \tilde{V}_l$ は $\tilde{V}_k, \tilde{V}_l$ の中の点 u, v の選び方にはよらないことに注意すべきである．このようにして作られた $V/\leftrightsquigarrow$ の上の関係(強連結成分の間の関係) $\rightsquigarrow$ は，今度は，半順序になる．

一般の有向グラフ G に対しては，G の **Hertz**(ヘルツ)**グラフ**と呼ばれるグラフ $\tilde{G}$ が次のように作られる．

- $\tilde{G}$ の点の集合 $\tilde{V}$ は G の強連結成分の集合 $\tilde{V} = \{\tilde{V}_1, \cdots, \tilde{V}_S\}$である．
- $\partial^+ e \in \tilde{V}_k$，$\partial^- e \in \tilde{V}_l$ であるような G の辺 $e(\in E)$が存在するとき $\tilde{\partial}^+ \tilde{e} = \tilde{V}_k, \tilde{\partial}^- \tilde{e} = \tilde{V}_l$ であるような $\tilde{G}$ の辺 $\tilde{e}(\in \tilde{E})$を設ける．

Hertz グラフはその構成法からして DAG であることは容易に確かめられる．

A.2 無向グラフ

有向グラフ $G = (V, E, \partial^+, \partial^-)$において，ある辺 $e(\in E)$を，$\partial^+ e = \partial^- e'$，$\partial^- e = \partial^+ e'$ であるような e' でおきかえることを，辺 e の**向きの反転**(reorientation)という．向きの反転は同時に複数本の辺に対しておこなうこともできる．そうすると，一つの有向グラフから辺の向きの反転によって多くの(正確には $2^{|E|}$ 個の)有向グラフの族が作られるが，それらに共通した性質にのみ着目するとき，その有向グラフの族のことを**無向グラフ**(undirected graph，unoriented graph)と呼

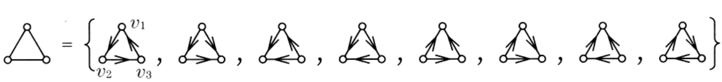


図 A.2.1 無向グラフ

ぶ．直観的には，有向グラフの辺の向きを矢印で表しておいて，その矢印を消したものを無向グラフだと思えばよい．

たとえば，有向グラフ G の二つの点 $u, v(\in V)$ の間に一方を始点とし他方を終点とするような道(有向道とは限らない)が存在するかどうかは，辺の向きの反転にはよらない性質である．u, v の間に道が存在することを $u \sim v$ と記す(そして $u \sim u$ と約束する)と，関係 $\sim$ は V の上の同値関係になる．$\sim$ に関する同値類を**連結成分**(connected components)という．連結成分の個数が 1 であるような，すなわちすべての 2 点の間に道が存在するような，グラフを**連結**(connected)であるという．すなわち，グラフの連結性は無向グラフの性質である．

点の集合 V が二つの部分 V_+ と V_- に分かれ($V = V_+ \cup V_-, V_+ \cap V_- = \emptyset$)，すべての辺 $e(\in E)$ が $\partial^+ e \in V_+$，$\partial^- e \in V_-$ であるような，すなわち V_+ の点同士，V_- の点同士を結ぶ辺がないようなグラフのことを**2 部グラフ**(bipartite graph)と呼ぶ．2 部グラフにおいては，点の二つの部分集合 V_+, V_- が指定されているから，辺の向きはそれから自然に定まるので，2 部グラフはとりあえずは無向グラフであるとみなしておいてよい．

無向グラフの辺の部分集合 $M(\subseteqq E)$ は，M の辺同士が端点を共有しないとき，**マッチング**(matching)と呼ばれる．与えられたグラフのマッチング M の中で $|M|$ が最大であるようなものを**最大マッチング**(maximum matching)，すべての点が M の辺のどれかの端点となっているようなものを**完全マッチング**(complete matching)などと呼ぶ．どんなグラフにも最大マッチングは当然存在するが，完全マッチングは存在するとは限らない．M がマッチングなら $2|M| \leqq |V|$，完全マッチングなら $2|M| = |V|$ であることは明らか．もちろん，完全マッチングなら最大マッチングである．

二つの無向グラフ G_1，G_2 が同一の点の集合 $V = V_1 = V_2$ ($|V|$ は偶数とする)をもち，辺の集合 E_1，E_2 には共通部分がない($E_1 \cap E_2 = \emptyset$)場合を考える．このとき，$G_1 \Cup G_2 = G$ は，点の集合 V と辺の集合 $E = E_1 \cup E_2$ をもつ無向グラフであるとする(辺の端点は G_1, G_2 におけるものとして自然に定義される)．G_1 の完全マッチング $M_1(\subseteqq E_1)$，G_2 の完全マッチング $M_2(\subseteqq E_2)$ に対して，次のような G における**交互サイクルの組**(set of alternating cycles)を作ることができる．すなわち，まず点番号の最も若い点 $v_0(\in V)$ は M_1 のどれかの辺 e_1 一つだけの端点であるから，e_1 の他方の端点 v_1 が定まる．すると，v_1 を端点とする M_2 の辺 e_2

がただ一つ定まり，e_2 の他の端点 v_2 が定まる．次に v_2 を端点とする M_1 の辺の他方の端点を v_3 とし，…と進めて行くと，どこかで $v_{l_1}=v_0$ となるが，このとき v_{l_1-1} と v_{l_1} を端点とする辺は M_2 の辺である．(v_0 を端点とする M_1 の辺は e_1 しかないから．) すなわち，v_0 から出発して $\{v_0, e_1, v_1, e_2, \cdots, e_{l_1}, v_{l_1}=v_0\}$ という M_1 ($\subseteqq E_1$)の辺と $M_2(\subseteqq E_2)$の辺とが交互に現れるような $G_1 \Cup G_2$ 上の"偶数長の"サイクル (l_1=偶数)が得られる．(サイクルや辺の向きはサイクルに含まれる辺がみな順方向に含まれるように定める．)このサイクルがすべての点を通っていなければ，残りの点の中で点番号が最小のもの(そしてそれは M_1 の辺のどれかの端点になっている)から出発して同様のことを繰り返す．M_1 も M_2 もマッチングであるから，同一点が再度現れることはなく，結局，何個かの互いに素な偶数長のサイクルがすべての点を覆うことになる．

逆に，$G_1 \Cup G_2$ 上の互いに素な交互サイクル(E_1 の辺と E_2 の辺が交互に並ぶサイクル)でその長さの総計が $n=|V|$(=偶数)である(したがってすべての点を覆う)ものは，G_1 上の完全マッチング M_1 と G_2 上の完全マッチング M_2 に分けられる．

この事実は，マッチングの簡単な性質であるが，いろいろな場面で重要な役割を果たす．

A.3　置換，順列とその符号

n 個のもの($1,2,\cdots,n$ と番号で指定することにする)を番号順に並べた並び方を基準にして，その並び方を$(\pi_1, \pi_2, \cdots, \pi_n)$と並べ替えたとする．$\pi_i$ は $1,\cdots,n$ のどれかに等しく，$i \neq j$ なら $\pi_i \neq \pi_j$ である．この並べ替えを**置換**(permutation)という．これは，集合 $V=\{1,\cdots,n\}$のそれ自身の上への**全単射**(bijection) $\pi: V \to V$ ($\pi_i=\pi(i)$)であるともみなせる．$(\pi_1,\cdots,\pi_n)$を$\{1,\cdots,n\}$の一つの並び方であるとみて**順列**と呼ぶこともある．(英語などではこれも permutation という．) 置換 π を

$$\pi=\begin{pmatrix} 1 & 2 & \cdots & n \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{pmatrix} \tag{A.3.1}$$

のように表記することもある．置換 π と π' の積は，写像としての合成写像 $\pi' \circ \pi$ であると定義する：

$$(\pi\bullet\pi')(i)=(\pi'\circ\pi)(i)=\pi'(\pi(i)). \tag{A.3.2}$$

この積に関して，置換の全体は一つの群をなす．(**対称群**(symmetric group)と呼ばれる．位数は $n!$．)

置換の最も簡単なもの，すなわち i_0 と j_0 ($i_0 \neq j_0$)を入れ換えるだけのもの：

$$\begin{aligned}&\pi(i)=i \qquad (i\neq i_0,j_0),\\ &\pi(i_0)=j_0,\quad \pi(j_0)=i_0\end{aligned} \tag{A.3.3}$$

のことを**互換**(transposition)と呼び，$(i_0,j_0)\,(=(j_0,i_0))$と略記する．

V の要素を点 $v_1,\cdots,v_n$ とみなし，置換 π に対して

$$\pi(i)=j \text{ のとき } \partial^+e_i=v_i,\quad \partial^-e_i=v_j \tag{A.3.4}$$

であるような辺 e_i を設けることにすれば，π を表現する有向グラフ G_π が作られる．G_π においては

$$\text{すべての } v\in V \text{ に対して } \mathrm{d}^+v=\mathrm{d}^-v=1 \tag{A.3.5}$$

であることは明らかであろう．また辺の総数 $|E|$ は点の総数 $|V|=n$ に等しい．

置換を表現するグラフに対しても自然に積が定義される：

$$G_{\pi'}\circ G_\pi=G_{\pi'\circ\pi} \tag{A.3.6}$$

置換を表現する有向グラフもこの積に関して一つの群を作るが，このとき群の単位元(すなわち**恒等置換**)を表す有向グラフは，すべての点にループが付いたグラフであり，π の逆元(すなわち逆写像) π^{-1} を表現する有向グラフ $G_{\pi^{-1}}$ は，G_π のすべての辺の向きを反転したものである．

G_π の上では，どこか任意の点 $v(1)$ から出発して，その点を始点とする辺(これは(A.3.5)により一意的に定まる)を通ってその辺の終点 $v(2)$ に至り，$v(2)$ を始点とする辺(これも一意的に定まる)を通ってその終点 $v(3)$ に至る，…というように続けて，一つの有向道を作って行くと，(A.3.5)によりいつかは出発点 $v(1)$ に戻る．このようにして得られる有向閉路$(v(1),\cdots,v(l))$は初等的である，すなわち一つのサイクル $C=(v(1),\cdots,v(l))$をなす．C がすべての点を通っていればそれで終わり，そうでなければ，このサイクルを $C_1=(v(1,1),\cdots,v(1,l_1))$と記して，$C_1$ に含まれない任意の点 $v(2,1)$ を選び同様のことを繰り返すと C_1 とは互いに素なサイクル $C_2=(v(2,1),\cdots,v(2,l_2))$が得られる．以下同様にして，$G_\pi$ はいくつかの互いに素なサイクル $C_1,\cdots,C_t$ (それぞれ長さが $l_1,\cdots,l_t$)から構成されることがわかる．(C_i の長さが 1，すなわち $C_i=(v(i,1))$ ならループ，長さが 2 なら，すなわち $C_i=(v(i,1),v(i,2))$ なら互換であること

にも注意.)

$$\sum_{i=1}^{t}|C_i|=\sum_{i=1}^{t}l_i=n \tag{A.3.7}$$

であることも明らかであろう.

このようなサイクルへの分解は実質的には一意であるが，さらに表現の曖昧性も除くために，G_π の互いに素なサイクルへの分解については，

(i)　各サイクル C_i の先頭は C_i に含まれる点の中で番号が最小のものであるとする：

$$v(i,1)<v(i,j) \qquad (j=2,\cdots,l_i), \tag{A.3.8}$$

(ii)　サイクル $C_1,\cdots,C_t$ はそれぞれの先頭の点の番号が昇順になるように並べる：

$$v(1,1)<v(2,1)<\cdots<v(t,1) \tag{A.3.9}$$

という条件を満たすように表すこととする.（その結果 $v(1,1)=1$ となることにも注意.）このように表された“置換の互いに素なサイクルへの分解”のことを，今後，DCD(disjoint cycle decomposition)と略称することとする.

$C_1,\cdots,C_t$ が互いに素であることから，各 C_i ごとに C_i に含まれない点は不動とした置換(**巡回置換**)のことも C_i と記す．すなわち，C_i に含まれない点にはみなループをつけた有向グラフを G_{C_i} と記せば，

$$\begin{aligned}\pi&=C_1\circ C_2\circ\cdots\circ C_t,\\ G_\pi&=G_{C_1}\circ G_{C_2}\circ\cdots\circ G_{C_t}\end{aligned} \tag{A.3.10}$$

である.(A.3.10)の表現では，積の順序はどう取ってもよいことに注意($C_1,\cdots,C_t$ が互いに素であるから).

次に，長さ l (≧3)の一つのサイクル $C=(v(1),v(2),\cdots,v(l))$ に着目しよう．それが

$$C=(v(1),v(2))\circ(v(2),v(3),\cdots,v(l)) \tag{A.3.11}$$

のように，一つの互換と長さが $l-1$ のサイクルとの積の形に表すことができることは容易に確められる.（今度は，二つのサイクルが互いに素ではない――$v(2)$ を共有する――こと，および積の順序にも注意！）(A.3.11)をさらに続行すれば，長さ l のサイクル C は

$$C=(v(1),v(2))\circ(v(2),v(3))\circ\cdots\circ(v(l-1),v(l)) \tag{A.3.12}$$

のように $l-1$ 個の互換の積として表すことができる.（$l=2$ ならはじめから

(A.3.12)の形になっている.)これを各 C_i についておこなえば，任意の置換 π は，(A.3.7)も見ると，

$$\sum_{i=1}^{t}(l_i-1)=n-t \tag{A.3.13}$$

個の互換の積として表すことができることがわかる.(C_i がループのときには $l_i-1=0$ で，互換は不要であるが，そのときも(A.3.13)は成立している.)

置換の互いに素なサイクルへの分解は一意に定まったが，互換の積としての表現は一意には定まらない. たとえば，

$$\begin{aligned}(1,2,3)&=(1,2)\circ(2,3)=(1,3)\circ(1,2)\\&=(2,3)\circ(1,2)\circ(1,3)\circ(2,3)\end{aligned} \tag{A.3.14}$$

のように. しかし，置換をどのような互換の積に表そうとも，そのときに使われる互換の総数の**偶奇**(parity)は，すぐ下に示すように，一意に確定するという事実がある. そこで，偶数個の互換の積として表されるような置換を**偶置換**(even permutation)，奇数個の互換の積として表されるような置換を**奇置換**(odd permutation)と呼ぶことができる. 置換の偶奇(これを$+1$，-1 と符号で特徴づけて，置換の**符号**(sign)と呼ぶこともある)を特徴づける上記の量(A.3.13)の $n-t$ に対して

$$n-t\equiv[\text{DCD における偶数長のサイクルの個数}] \bmod 2 \tag{A.3.15}$$

であることにまず注意しよう. 実際，長さ l_i が奇数であるようなサイクルについては

$$l_i-1\equiv 0 \mod 2 \tag{A.3.16}$$

であるから，(A.3.13)を mod 2 で見れば(A.3.15)が成立することがわかる.

(A.3.15)によって置換の偶奇を特徴づけておけば，置換を互換の積として表す表し方にはよらない特徴づけになるが，実際に(A.3.15)が互換によって符号が反転することを確認しておく必要があろう. いま置換 π (あるいはそれを表現するグラフ G_π)と互換(i_0, j_0)との積である置換

$$\pi'=(i_0, j_0)\circ\pi \tag{A.3.17}$$

を考える.($\pi'=\pi\circ(i_0, j_0)$としても議論はほぼ同様.) π の DCD(A.3.10)に対して，点 i_0 と j_0 ($i_0<j_0$ としてよい)は

(a)　両方とも一つのサイクル C_i に属す

か

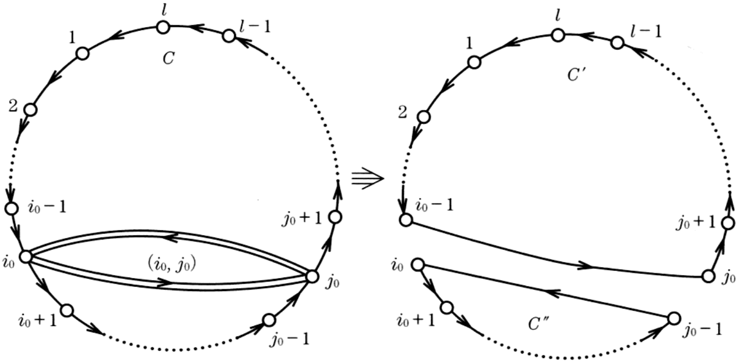


図 A.3.1　$(i_0, j_0)\circ C = C'\circ C''$

(b)　i_0 が C_i に属して j_0 が C_j に属す$(i \neq j)$

かのどちらかである．

π の DCD が互換(i_0, j_0)との積を作ることによって変化するのは，(a)の場合には C_i においてだけ，(b)の場合には C_i と C_j においてだけであるから，それを調べてみよう．

まず，(a)の場合には，C_i を簡単のため $C=(1,2,\cdots,l)$ と書いておくと，図 A.3.1 からもわかるように，

$$\begin{aligned}&(i_0, j_0)\circ(1,2,\cdots,l)\\&\quad=(1,\cdots,i_0-1,j_0,j_0+1,\cdots,l)\circ(i_0,i_0+1,\cdots,j_0-1)\end{aligned} \tag{A.3.18}$$

であるから，C は

$$C'=(1,\cdots,i_0-1,j_0,j_0+1,\cdots,l) \tag{A.3.19}$$

と

$$C''=(i_0,i_0+1,\cdots,j_0-1) \tag{A.3.20}$$

に分かれるが，C' と C'' に含まれる点の総数は不変であるから

$$l=l'+l'' \tag{A.3.21}$$

である(l', l'' は C', C'' の長さ)．すなわち，π' の DCD は

$$\pi'=C_1\circ C_2\circ\cdots\circ C_{i-1}\circ C_i'\circ C_i''\circ C_{i+1}\circ\cdots\circ C_t \tag{A.3.22}$$

で，点の総数が不変でサイクルの総数 t が 1 だけ増えることがわかる．このようにして，(A.3.15)の偶奇は反転する．

同様にして，(b)の場合には，π の DCD の C_i と C_j が合併して π' においては

一つのサイクルになることが見られ，やはり(A.3.15)の偶奇は反転する．

このようにして，一つの置換を互換の積として表したとき，そこに使われる互換の総数の偶奇は，その表し方には無関係であることがわかる．

以上をまとめると，置換 π の偶奇(あるいは符号)は，次の3通りの等価な定義が可能である．

(i) π を互換の積として表したとき，そこに現れる互換の総数の偶奇．

(ii) π を表現する有向グラフ G_π の点の数を n，G_π を構成する互いに素なサイクルの個数を t としたとき，$n-t$ の偶奇．

(iii) G_π を構成する互いに素なサイクルのうち，長さが偶数であるものの個数の偶奇．

ここでは詳しくは立ち入らないが，置換を $\pi=\begin{pmatrix} 1 & 2 & \cdots & n \\ \pi_1 & \pi_2 & \cdots & \pi_n \end{pmatrix}$ と表したとき，

(iv) $i,j\ (=1,\cdots,n)$ に対して $i<j$ かつ $\pi_i>\pi_j$ であるような対(i,j)(これを**反転**(inversion)という)の個数の偶奇

も，上記(i)，(ii)，(iii)と等価であることを示すのはそう難しいことではない．これとほとんど同じことであるが，n 個の変数 $x_1,\cdots,x_n$ の基本交代式(あるいは差積)

$$P(x_1,\cdots,x_n)=\prod_{\substack{i,j=1,\cdots,n \\ i<j}}(x_i-x_j) \tag{A.3.23}$$

(これまでの章では $\pi(x_1,\cdots,x_n)$ などと書くことが多かったが，ここでは置換 π と区別するため π という記号を避けた)に対して

(v) $$P(x_{\pi_1},\cdots,x_{\pi_n})/P(x_1,\cdots,x_n)=\begin{cases} 1 \text{ なら } \pi \text{ は偶置換,} \\ -1 \text{ なら } \pi \text{ は奇置換} \end{cases} \tag{A.3.24}$$

と定義することもできる(2.16.1項など参照)．

どの定義に従っても同じことであるが，置換の積の偶奇については次のことはほとんど自明である．(定義(i)によるのが一番簡単であろう．)

(a) 偶置換と偶置換の積は偶置換である．

(b) 偶置換と奇置換の積は奇置換である．

(c) 奇置換と偶置換の積は奇置換である．

(d) 奇置換と奇置換の積は偶置換である．

2.5.1項で導入した記号 $\varepsilon^{i_1\cdots i_r}_{j_1\cdots j_r}$ の正当性も上のようにして裏付けられる．

A.4　除算を使わない行列式の計算法

行列式の定義(2.6.1)，(2.6.2)は，n 次正方行列 $A=[a_{ij}]$の要素 a_{ij} の積と和だけを用いて表されている．したがって，a_{ij} が体に属さなくても，少なくとも可換な環に属してさえいれば定義には意味があるはずである．しかし，なるべく少数回の四則演算を用いて行列式の値を計算しようとの努力の結果，現在 $O(n^3)$ の演算回数の算法が“実用的”であるとして広く受け入れられている[1)]．

不思議なことに，($O(n^\alpha)$，$\alpha<3$ の算法はもちろん)$O(n^3)$の算法においては除算が重要な役割を果たしている．そこで，除算抜きで，すなわち加減算と乗算のみで行列式の値を計算する多項式オーダーの算法はないかという疑問は当然生じる．実は，古くから知られているものとして演算回数が $O(n^4)$ の Leverrier-Faddeev 法（注意 5.3.3)などがあった．((5.3.13)には見掛け上“k で割る”部分があるが，これは(5.3.12)からも明らかなように，同類項が k 個あるものをまとめて一つにするためのものであるから，本質的な除算ではない．c_{n-k} 自身が行列 A の主小行列式の和であることに注意.)

式(2.6.1)，(2.6.2)を直接そのまま計算したのでは演算回数は $O(n!)$ になるし，Laplace 展開を用いたり 2 分法を用いたりしても演算回数は多項式オーダーには収まらない．ところがここ十数年くらいの間に M. Mahajan，V. Vinay，その他の人達によって，無駄な項を加えたり引いたりするだけで，(2.6.1)などを直接計算するのに等価ではあるが演算回数が $O(n^4)$ の算法が工夫された．ここではまずその紹介をしよう[2)]．(Pfaffian については A.5 節で紹介する.)

n 次正方行列 $A=[a_{ij}]$に対して，辺に**重み**(weight)のついた有向グラフ $G(A)$を次のように作る．

- $G(A)$の点の集合は $V=\{1,2,\cdots,n\}$である；
- $G(A)$は，$a_{ij}\neq 0$ のときそしてそのときに限り，$\partial^+ e_k=i$，$\partial^- e_k=j$ であるよ

1)　$O(n^\alpha)$ ($\alpha\doteqdot 2.3$)のような算法も理論的には可能であることが知られているが，行列の疎性が利用できないとか，cn^α の係数部分の c が途方もなく大きくなるとかいうことにより，実際家達からは見向きもされていないようである．

2)　主として Meena Mahajan and V. Vinay：Determinant：Combinatorics, Algorithms, and Complexity. *Chicago Journal of Theoretical Computer Science*, Vol. 1997, Article 5 (December 1997), pp. 1-26，による．

うな(すなわち点 i を始点とし点 j を終点とする)辺 e_k をもつ．平行辺は考えないから $e_k=(i,j)$ と記して混乱は起きない；

・辺(i,j)の重みは $w(i,j)=a_{ij}$ である．

$G(A)$は，一般に，多くの辺部分グラフをもつが，その中で置換 π を表すもの，すなわち

$$\mathrm{d}^+i=\mathrm{d}^-i=1 \qquad (i=1,\cdots,n) \tag{A.4.1}$$

であるような辺部分グラフ $G_\pi(A)$ とその DCD 表現に着目する．このとき $G_\pi(A)$を経由しなくとも，$G(A)$の“互いに素なサイクルによる点集合 V の**被覆**(disjoint cycle cover；DCC と略記)”を考えることができる．ある DCC を $\{C_1,C_2,\cdots,C_t\}$とすれば，それが「すべての点を通る」ことと「各サイクルの長さの合計が $\sum_{i=1}^{t} l_i=|V|=n$ である」こととは同じことである．そこである DCC を構成するサイクルが $C_1,\cdots,C_t$ であるとき，その DCC の重みを

$$w(\mathrm{DCC})=C_1,\cdots,C_t \text{ に含まれる辺の重みの積}, \tag{A.4.2}$$

その符号を

$$\varepsilon(\mathrm{DCC})=(-1)^{n-t}=\begin{cases}+1, & \text{DCC が偶置換を表すとき},\\ -1, & \text{DCC が奇置換を表すとき}\end{cases} \tag{A.4.3}$$

と定めることにすれば，行列 A の行列式が

$$\det A=\sum_{G(A)\text{のすべての DCC}} \varepsilon(\mathrm{DCC})\,w(\mathrm{DCC}) \tag{A.4.4}$$

と表されることは，行列式の定義(2.6.1)の言い換えにすぎない．

注意 A.4.1　行列式を DCC で表現する式(A.4.4)は，DCC の定義を少し修正して，“互いに素なサイクルで，それらの長さの合計が m $(\leqq n)$に等しいもの”(これを大きさ m の“部分 DCC”と呼び DCC_m と記すことにする)および“それにより被覆される点の部分集合 U $(\subseteqq V)$ $(|U|=m)$”を考えることにより，

$$c_{n-m}=\sum_{\substack{G(A)\text{の上の大きさ }m\\ \text{のすべての DCC}_m}} \varepsilon(\mathrm{DCC}_m)\,w(\mathrm{DCC}_m) \tag{A.4.5}$$

という形で，行列 A の特性多項式の係数 c_{n-m} を表すことができる．ここで，ε と w は，$\mathrm{DCC}_m=\{C_1,\cdots,C_t\}$ $(\sum_{i=1}^{t} l_i=m)$に対して，それぞれ

$$\begin{aligned} w(\mathrm{DCC}_m)&=C_1,\cdots,C_t \text{ に含まれる辺の重みの積},\\ \varepsilon(\mathrm{DCC}_m)&=(-1)^{m-t}\end{aligned} \tag{A.4.6}$$

であるとする．実際，この DCC_m が被覆する点の部分集合 U は，$C_1,\cdots,C_t$ が互いに素であるから，$|U|=m$ であるが，同一の U を被覆する DCC_m だけを考えれば，(A.4.5)の和の中でそれらは $G(A)$の点部分グラフ $G_U(A)$に対応する行列式，

すなわち A の行と列が U の点に対応する A の m 次の主小行列式の(A.4.4)の形の表現になっている．(A.4.5)はそれらをすべての大きさ m の U についてさらに和を取ったものになっている．これは本文の注意 5.1.5 の式(5.1.13)により c_{n-m} に等しい．

$G(A)$が**完全グラフ**(complete graph)，すなわち $a_{ij} \neq 0$ $(i,j=1,\cdots,n)$のときには，置換を表すような辺部分グラフは $n!$ 個あるから，(A.4.4)の $\sum$ は $n!$ 個の項をもち，各項は n 個の $w(i,j)$の積であるから，(A.4.4)をこのまま計算したのでは $n!(n-1)$回の乗算と $n!$ (うるさくいえば $n!-1$ かも)回の加算(ε が -1 のときは減算)が必要となる．

Mahajan，Vinay らは，サイクルを一旦有向閉路に一般化することを考えた．すなわち，**初等的とは限らない**有向閉路

$$C_i=(v(i,1),\ v(i,2),\ \cdots,\ v(i,l_i)) \tag{A.4.7}$$

で，$l_i \geqq 2$ なら

$$v(i,1)<v(i,j) \qquad (j=2,\cdots,l_i) \tag{A.4.8}$$

を満たすものを考えた．($l_i=1$ のときには $C_i=(v(i,1))$はループで，条件(A.4.8)は無効.) (A.4.8)は，有向閉路 C_i の先頭 $v(i,1)$が閉路に含まれる点の中で最も番号が小さく，かつ C_i の中には $v(i,1)$は 1 回しか含まれないことを意味している．

さらに，t 個の有向閉路 C_1，C_2，…，C_t を，

- 有向閉路に含まれる辺の延べ数(複数回含まれるときにはその回数だけ重複して数える)が n に等しい；
- 各 C_i の先頭の点の番号 $v(i,1)$は(A.4.8)を満たし，かつ異なる C_i の間では狭義単調増大である，すなわち

$$v(1,1)<v(2,1)<\cdots<v(t,1) \tag{A.4.9}$$

を満たすように選ぶ(かなり技巧的ではあるが)．このように選ばれた有向閉路の組を単に**有向閉路列**(closed walk sequence)と呼び，Mahajan，Vinay に従って ClWS と略記することにする．

ClWS では，それを構成する有向閉路の先頭の点番号が重複して現れないこと以外は，同一番号の点があちこちに重複して現れてもよいとしているところが DCC と異なるところである．しかし，DCC は ClWS の特殊な場合になっている

ことに注意しよう．

このような ClWS の重みは，ClWS＝$(C_1, \cdots, C_t)$ のとき，

w(ClWS)＝$C_1, \cdots, C_t$ に含まれる辺の重みの積

（複数回現れる辺の重みはその回数だけ重複して積を取る），　(A.4.10)

その符号は

$$\varepsilon(\mathrm{ClWS}) = (-1)^{n-t} \tag{A.4.11}$$

であるとする．(DCC のときと形式が揃えてあることに注意．)

(A.4.8)，(A.4.9)の条件だけでは，$v(1,1)=1$ となるとは限らないが，さらに $v(1,1)=1$ という条件を付け足しても $\det A$ の計算については差し支えない．点や辺が ClWS の中に重複して現れてもよいということは，点が含まれる延べ回数が n であるという制限により，ClWS の中に現れない点があってよいということにもなることに注意されたい．

次に(A.4.4)と形式的に同じ

$$\det' A = \sum_{\substack{G(A)\text{上の}\\ \text{すべての ClWS}}} \varepsilon(\mathrm{ClWS})\, w(\mathrm{ClWS}) \tag{A.4.12}$$

を考える．(A.4.12)の右辺の項の中には(A.4.4)の右辺の項がすべてその一部として含まれているが，それら以外の項が互いに打ち消し合うことが以下のように示されるので結局 $\det' A = \det A$ となるということが興味あるところである．さらに(A.4.12)の $\det' A$ の計算が DAG を利用することにより多項式オーダーで実行できるということも，もう一つ興味あるところである．

いま，ある ClWS＝$(C_1, \cdots, C_t)$ が DCC でなかったとする．すると，C_t, $C_{t-1}, \cdots$ の順に調べて行くと「$C_t, \cdots, C_{k+1}$ まではそれぞれの C_i がサイクルであってかつ互いに素であるが，C_k を取り込むとそうではなくなる」というような k ($\geqq 1$) が存在する．(C_t がサイクルでないときには $k=t$ とする．この ClWS が DCC ならば $k=0$ であるとする．) このとき次の二つの可能性(a)，(b)がある．

(a)　$C_k = (v(k,1), v(k,2), \cdots, v(k,l_k))$ の点を先頭から順に調べて行くと「$v(k,m)$ がはじめて $C_{k+1}, \cdots, C_t$ のどれかの点と一致する」というような m (>1) ($m=1$ は(A.4.8)，(A.4.9)によりありえない)に行き当たる．

(b)　C_k がサイクルでない(初等的でない)ことがわかる；すなわち，C_k の点を先頭から順に調べて行くと「はじめて，$v(k,m) = v(k,m-l)$ となる」よ

うな $m\ (>1)$，$l\ (\geqq 1)$ に行き当たる．($m>1$ は(A.4.8)による．また，この場合には，$v(k,m-l),\cdots,v(k,m-1)$ の中には同一の番号は 2 度以上現れないから，$(v(k,m-l),v(k,m-l+1),\cdots,v(k,m-1))$ は初等的であり，したがって長さ l のサイクルである．)

上の(a)と(b)が同時に，すなわち同じ m に対して，生起することはないことに注意しよう．なぜならば，ある m に対して(a)と(b)が成り立てば，(b)により $m'=m-l$ に対しても(a)が成り立っていたことになり，“はじめて”ということに反するからである．

(b)が生起する場合：——サイクル$(v(k,m)=v(k,m-l),v(k,m-l+1),\cdots,v(k,m-1))$の点の中で番号の最小のものが先頭に来るように巡回させたものを C' とおき，

$$C'_k=(v(k,1),\cdots,v(k,m-l),v(k,m+1),\cdots,v(k,l_k)) \qquad \text{(A.4.13)}$$

とおく．C'_k の先頭は C_k のそれに等しく，その長さは C_k の長さより l だけ小さい．$\mathrm{ClWS}(C_1,\cdots,C_t)$ が条件(A.4.8)，(A.4.9)を満たし，(b)が生起しているのであるから，(A.4.9)が満たされるように C' を$(C_1,\cdots,C_{k-1},C'_k,C_{k+1},\cdots,C_t)$のどこか，それも $C_{k+1},\cdots,C_t$ のどこかに挿入して新しい

$$\mathrm{ClWS}'=(C_1,\cdots,C_{k-1},C'_k,C_{k+1},\cdots,C',\cdots,C_t) \qquad \text{(A.4.14)}$$

を作ることができる．ClWS′ もそれが含む辺の数は，重複も数えて，もとの ClWS と等しいから，$w(\mathrm{ClWS}')=w(\mathrm{ClWS})$ であり，t が 1 だけ増加するから $\varepsilon(\mathrm{ClWS}')=-\varepsilon(\mathrm{ClWS})$ である．このようにして，ClWS と ClWS′ とは(A.4.12)の和の中で互いに打ち消し合う．

(a)が生起する場合：——$v(k,m)$ と一致する点を含む $C_{k+1},\cdots,C_t$ のどれか（これは $C_{k+1},\cdots,C_t$ が互いに素であるという仮定から一意に確定する）を $C_j\ (k<j\leqq t)$ とし，$C_k=(v(k,1),\cdots,v(k,m),\cdots,v(k,l_k))$ の $v(k,m)$ の直後にサイクル C_j を($v(k,m)$が先頭に来るように巡回させた後)挿入して出来る有向閉路を C'_k とする．C'_k の先頭の点の番号は C_k のそれと等しく，また(A.4.8)，(A.4.9)により $v(k,1)<v(k,m)$ かつ $v(k,1)<v(j,1)$ で C_j のすべての点の番号(そしてそれは C_j を巡回させても同じ)は $v(j,1)$ より小さくないから，C'_k に対しても(A.4.8)が成り立つ．そこで新しい

$$\mathrm{ClWS}'=(C_1,\cdots,C_{k-1},C'_k,C_{k+1},\cdots,C_{j-1},C_{j+1},\cdots,C_t) \qquad \text{(A.4.15)}$$

を作れば，$w(\mathrm{ClWS}')=w(\mathrm{ClWS})$，$\varepsilon(\mathrm{ClWS}')=-\varepsilon(\mathrm{ClWS})$となる．(A.4.12)の中ではこれら二つの ClWS と ClWS′ の項は打ち消し合う．

上記の(a)，(b)に基づいた「一つの ClWS からもう一つの ClWS′ を作る」方法によって $G(A)$上のすべての ClWS の上の一つの写像 φ が定義される．φ^2 が恒等写像であることも容易にわかる．(DCC は φ によって自分自身に写されるとみなす．) このような φ は**対合**(involution)と呼ばれる．この φ によって(A.4.12)の右辺で DCC 以外の ClWS は必ず重みが等しくて符号が反対になるものと対をなしているので，$\sum$ を取ると打ち消し合って(A.4.4)の形の和だけが残るという話になる．

このような一見“ややこしい”議論をしたのは，(A.4.12)の形の和の計算には多項式オーダーの算法が使えるからである．そのために，(A.4.8)，(A.4.9)の条件を考慮しながら，(A.4.12)の各項の ClWS を同時並行的に作り出して行くことを考える．各項 $w(\mathrm{ClWS})$ は辺の重みの積の形になっているが，この ClWS を構成する有向閉路を(A.4.8)，(A.4.9)の条件を満たすように左から順に部分的に作って行き，$m-1$ 個の辺の重みの積(符号付き)まで作られたとしよう．このとき，いくつかの有向閉路は完成し，新しい未完成の有向閉路(ということはその先頭部分の有向道)が作られる途中であるとする．この状態は m に関して“層化”して考えられる．一般に，第 m 層の状態は

現在作られつつある有向閉路の先頭点の番号 i

と

伸びつつある有向道の最終点の番号 j

とによって表現できる．この場合条件(A.4.8)により $i<j$ であるが，特別な場合として，それまで伸ばしてきた有向道(先頭点は i' であるとする)が有向閉路として完成し，新しい有向閉路が作られ始める状態を表すには，条件(A.4.9)に鑑み $i=j$ $(i>i')$とする．

このように層化された状態と，m を1ずつ増加させるときの状態の変化は一つの DAG で表すことができる．

$m=1$ の層はすべての出発の状態であるから，状態$(m\,;\,i,j)=(1\,;\,i,i)$ $(i=1,\cdots,n)$からなる．

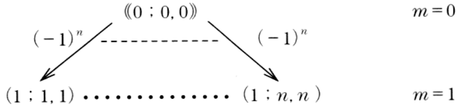


図 A.4.1　算法の開始

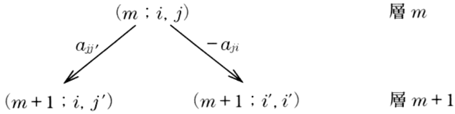


図 A.4.2　ClWS の成長

$m=0$ の状態も考えて，状態$(0;0,0)$を設定して，そこから $m=1$ の点すべてに $w(e)=(-1)^n$ なる重みをもつ辺 e を設ける．(この重み，符号は，後に述べるように，新しい有向閉路を出発させる$(-1)^{n-t}$ の t を1だけ増やすので，最初は$(-1)^n$ という符号をつけておくということである．)

第 m 層の状態$(m;i,j)$からは，$j'>i$ でかつ $a_{jj'}\neq 0$ なる第 $m+1$ 層のすべての状態$(m+1;i,j')$へ，重み $w(e)=a_{jj'}$ の辺を設け(有向道を伸ばすことに対応)，さらに $a_{ji}\neq 0$ ならば $i'>i$ なるすべての状態$(m+1;i',i')$へ重み $w(e)=-a_{ji}$ の辺 e を設ける．(有向閉路を完成させるために点 j から i に戻るとともに，それにより有向閉路の数が1だけ増加して符号が反転することを反映させて $-a_{ji}$ という重みをつける．$i'>i$ は条件(A.4.9)を満たすため．)

第 n 層ですべてが終了するから，そこまで伸びてきた有向道は有向閉路として完成させなければならないので，第 $n+1$ 層に状態$(n+1;0,0)$を設けて，第 n 層のすべての状態$(n;i,j)$から$(n+1;0,0)$へ重み $w(e)=-a_{ji}$ の辺 e を設ける(もちろん $a_{ji}\neq 0$ ならば)．そうすれば，このDAG上で第0層の1点から出発して第 $n+1$ 層の点に至るすべての有向道が，もとのグラフ $G(A)$上のClWSに対応することになり，その符号付き重みが(A.4.12)の右辺の各項 ε(ClWS) w(ClWS)に等しいことになる．

これで(A.4.12)がA.1節の(A.1.6)の形になっていることがわかった．そこで(A.1.6)の計算法にならって，各状態(すなわちDAGの点)$(m;i,j)$に，ポテンシャル $W(m;i,j)$を付随させ，以下の通りの計算をおこなうことによって $\det A$ の値が $W(n+1;0,0)$に作られる．

算法 A.4.1

[1]　すべての $i=1,\cdots,n$ に対して $W(1\,;\,i,i):=(-1)^n$. すべての $m=1,\cdots,n,n+1$ に対して，すべての存在する点 $(m\,;\,i,j)$ に対して，$W(m\,;\,i,j):=0$.

[2]　$m=1,2,\cdots,n$ に対して，すべての存在する点 $(m\,;\,i,j)$ と $(m+1\,;\,i',j')$ について，

[2.1]　$(m\,;\,i,j)$ が $(m+1\,;\,i,j')$ $(j'>i)$ に辺で結ばれているときには

$$W(m+1\,;\,i,j'):=W(m+1\,;\,i,j')+W(m\,;\,i,j)\,a_{jj'}\,;$$

[2.2]　$(m\,;\,i,j)$ が $(m+1\,;\,i',i')$ $(i'>i)$ に辺で結ばれているときには

$$W(m+1\,;\,i',i'):=W(m+1\,;\,i',i')-W(m\,;\,i,j)\,a_{ji}.$$

[3]　すべての存在する点 $(n\,;\,i,j)$ について

$$W(n+1\,;\,0,0):=W(n+1\,;\,0,0)-W(n\,;\,i,j)\,a_{ji}.$$

注意 A.4.2　(A.4.4) に対して (A.4.5) を考えたように，(A.4.12) に対しても"大きさ m の部分 ClWS"を考えることができる．すなわち，(A.4.7)，(A.4.8) を満たす有向閉路の列 $\{C_1,\cdots,C_t\}$ で，$C_1,\cdots,C_t$ に含まれる辺の数の合計(重複も数えて)が m に等しいようなものを ClWS_m と記すことにする．すると，DCC と ClWS との間の関係はそのまま DCC_m と ClWS_m との間に成り立つことがわかる．そこで $\det A$ の計算のために構成した DAG とその上での算法 A.4.1 をほんの少し加工することによって，行列 A の特性多項式の係数もすべて計算できるようになる．

具体的には，点 $(n+1\,;\,0,0)$ 以外に $(2\,;\,0,0),(3\,;\,0,0),\cdots,(n\,;\,0,0)$ を導入し，算法 A.4.1 の中に

[2.3]　すべての存在する点 $(m\,;\,i,j)$ に対して

$$W(m+1\,;\,0,0):=W(m+1\,;\,0,0)-W(m\,;\,i,j)\,a_{ji}$$

を追加すれば，$W(m+1\,;\,0,0)$ に特性多項式の係数 c_{n-m} が作られる．(そして，算法 A.4.1 の中の[3]は[2.3]に含まれるので，取り去る．)

注意 A.4.3　計算上大切なことであるが，DAG 上の計算であるので，$W(m\,;\,i,j)$ をすべての m について常に保持しておく必要はない．$W(m\,;\,i,j)$ を計算するためには $W(m-1\,;\,i,j)$ だけしか使われないから，$W(m\,;\,i,j)$ の計算が済んでし

まえば $W(m-1;i,j)$ は最早不要で，それに相当する記憶場所は $W(m+1;i,j)$ の計算に使うことができる．(算法 A.4.1 の[1]における初期化はその度ごとにおこなう．)各層 m ごとに状態(すなわち点)$(m;i,j)$ は約 $n^2/2$ 個あり，そこから**出る**辺は約 $n^3/3$ 個あることに注意．(計算用 DAG の構成法から明らか．) しかも，DAG の構成法からわかるように，辺はあらかじめ陽に描いておかなくても計算は進めることができる．このようにして，算法 A.4.1 の計算量は $O(n^4)$ であると見積もることができる．

例 A.4.1　算法 A.4.1 に従って 3 次の正方行列($n=3$)の行列式(およびその特性多項式の係数 c_1, c_2)を計算するときの DAG を，辺の重み(と符号)を付けて，図 A.4.3 に描いてある．そこで，

《 ; , 》は新しい有向閉路が作り始められる状態，あるいは行列式(あるいは主小行列式の和)の値を作り出すための状態，

┼┼┼┼┼┼┼ は DCC のサイクル列，

-------- は c_{n-m} の計算もするときの DAG 上の辺

を表す．DCC のみを計算するのは厄介であるが，ClWS にしてしまえば DAG

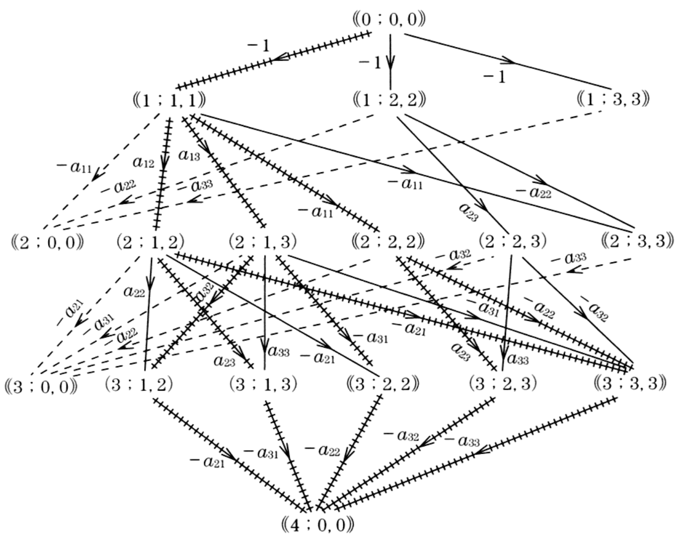


図 A.4.3

が素直に利用できること，また，DCC にならない ClWS は相互に打ち消し合っていることに注意されたい． □

注意 A.4.4　この計算法と，除算が必須の Gauss の消去法系統のものとを比較すると，もちろん前者の計算量は $O(n^4)$ で後者は $O(n^3)$ であるが，行列の疎性の影響はこの算法の方がはるかに見やすい．前者の計算量は，$O(G(A)$ の辺の数 $\times n^2)$ である．Gauss 流の算法であると，特殊な形の疎行列でないと計算途中でのいわゆる fill-in(すなわち，もとの行列で零であったところに非零要素が詰まってくる現象)のために，疎性の活用技術にすっきりした話がしづらい．

注意 A.4.5　例 A.4.1 を見てもわかるように，計算用の DAG の点(状態)の中には $\det A$ の計算には結局貢献しないものがある．(図 A.4.3 では，《1；3,3》と《2；3,3》，それに《1；2,2》と (2；2,3) も行列式の値には貢献していない．) 一般に第 m 層における《m；i,i》という状態では，それまでに完成した有向閉路の長さの合計は $m-1$ である――したがって，そこには $m-1$ 個以下の異なる番号の点しか現れない．一方，その後に現れる有向閉路は，先頭の番号が $i+1$ 以上 n 以下であるから(条件(A.4.8)，(A.4.9)！)，そこにはたかだか $n-i$ 個の異なる番号の点しか現れえない．そこで，このような有向閉路上にはたかだか $(m-1)+(n-i)$ 個の異なる番号の点しか含まれない．ところが，行列式の値の計算に最終的に貢献する有向閉路には n 個の点すべてが含まれていなければならないから，$(m-1)+(n-i)=(m-i)+(n-1)$ により $m \leqq i$ であると $(m-i)+(n-1) \leqq n-1$ となって n 個の点すべてが含まれるわけには行かない．そこで，DAG を構成するとき $m \leqq i$ であるような状態《m；i,i》を作らないようにすれば無駄な計算(あるいは操作)が省けることになる．

しかし，この簡単化だけでは DCC にならない ClWS をすべて切り落とすことはできないことにも注意されたい．しかも，このような簡単化をすると，特性方程式の係数が副産物として得られるという利点が失われることになる．

A.5　Pfaffian に対する組合せ論的接近法と除算を使わない Pfaffian の計算法

$|V|=n$(＝偶数)なる点の集合 V をもつ二つのグラフ G_1, G_2 の辺の集合を E_1, E_2 とし($E_1 \cap E_2=\emptyset$)，G_1 上の一つの完全マッチングを M_1，G_2 上の一つの完全マッチングを M_2 などと記す(A.2 節参照)．G_i $(i=1,2)$ は本来無向グラフではあるが，便宜的に辺 e は $\partial_i^+ e < \partial_i^- e$ であるように，すなわち各辺はその端点

の番号の小さい方から大きい方へと向きづけられているものとする．

二つの n 次(偶数次)反対称行列を A,B とし，G_1 に対しては $\partial_1^+ e=i$，$\partial_1^- e=j$ $(i<j)$ のとき $w(e)=a_{ij}$ として e の**重み**を定義する．G_2 に対しても，同様に，$\partial_2^+ e=i$，$\partial_2^- e=j$ $(i<j)$ のとき $w(e)=b_{ij}$ とする．

G_1 上の完全マッチング $M_1(\subseteqq E_1)$ に対しては，その重みを $w(M_1)=\prod_{e\in M_1} w(e)$ で，その符号を $\varepsilon(M_1)=\varepsilon^{1\ 2\ \cdots\ n-1\ n}_{i_1\ j_1\ \cdots\ i_{n/2}\ j_{n/2}}$ で定義する．ここで $M_1=\{e_1,e_2,\cdots,e_{n/2}\}$ に対し $\partial_1^+ e_1=i_1$，$\partial_1^- e_1=j_1$，…とする．$\varepsilon^{1\ 2\ \cdots\ n-1\ n}_{i_1\ j_1\ \cdots\ i_{n/2}\ j_{n/2}}$ は置換$\begin{pmatrix}1 & 2 & \cdots & n-1 & n\\ i_1 & j_1 & \cdots & i_{n/2} & j_{n/2}\end{pmatrix}$の符号である．この際，マッチング M_1 が定まればその符号に要素 $e_1,e_2,\cdots$の並び順および e_k の向きづけには無関係に $\varepsilon(M_1)\,w(M_1)$ は定まることに注意すべきである．実際，$e_1,e_2,\cdots$の並びを変えても$(i_1\ j_1\cdots i_{n/2}\ j_{n/2})$には偶数回の互換が作用するだけであるから $\varepsilon(M_1)$の符号は変化しない．またある e_k の向きづけ(ここでは一応それを便宜的に定めはしたが)を逆にすると $w(e_k)$の符号が反転する(A が反対称であることによる)と同時に $\varepsilon(M_1)$にも一つの互換が作用して符号が反転するからである．そこで，2.9 節での Pfaffian の定義は

$$\mathrm{Pf}\,A=\sum_{M_1:\,G_1\text{上の完全マッチング}} \varepsilon(M_1)\,w(M_1) \tag{A.5.1}$$

と書くことができる．ここで M_1 は集合として同じものは 1 回しか数えないこととする．Pf B，G_2，M_2 についても同様である．

しかるに，G_1 上の完全マッチング M_1 と G_2 上の完全マッチング M_2 に関しては A.2 節の最後で述べた通り，$M_1\cup M_2$ は $G_1 \Cup G_2$ 上のそれぞれ偶数長の互いに素ないくつかの交互サイクルからなる．逆に $G_1 \Cup G_2$ 上の互いに素な交互サイクル$\{C_1,\cdots,C_t\}$で長さの合計が n のものは，G_1 上の完全マッチング M_1 と G_2 上の完全マッチング M_2 に分けられる．すなわち，$M_1\cup M_2$ は $G_1 \Cup G_2$ の上の互いに素な交互サイクルによる被覆(これを DACC(disjoint alternating cycle cover)と略称する)DACC$=\{C_1,\cdots,C_t\}$と一対一に対応する．ここで，明らかに，

$$w(M_1)\cdot w(M_2)=w(\mathrm{DACC}) \tag{A.5.2}$$

である．

次に，$\varepsilon(M_1),\varepsilon(M_2),\varepsilon(\mathrm{DACC})$の関係を調べる．交互サイクル C_k の長さを l_k とし，それが通る点の番号を $v(k,i)$ $(i=1,\cdots,l_k)$とする．すなわち

$$C_k=(v(k,1),\cdots,v(k,l_k)) \tag{A.5.3}$$

(ただし，一般性を失わずに，$v(k,1)$ を $v(k,2)$ に結ぶのは G_1 の辺，すなわち M_1 の辺，であるとする)とすれば，C_k が交互サイクルであることにより

$$\varepsilon(M_1)=\varepsilon_{v(1,1)\,v(1,2)\cdots v(1,l_1)\,;\,v(2,1)\cdots v(2,l_2)\,;\,\cdots}^{\;1\quad\;\; 2\quad \cdots\quad l_1\;\;\,;\;\, l_1+1\,\cdots\, l_1+l_2\;;\;\cdots},$$

$$\varepsilon(M_2)=\varepsilon_{v(1,2)\,v(1,3)\cdots v(1,l_1)\,v(1,1)\,;\,v(2,2)\cdots}^{\;1\quad\;\; 2\quad\; \cdots\;\; l_1-1\quad l_1\;\;;\;\; l_1+1\,\cdots} \tag{A.5.4}$$

である．$\varepsilon(M_2)$ は

$$\varepsilon(M_2)=\varepsilon_{v(1,1)\,v(1,2)\,v(1,3)\cdots v(1,l_1)\,;\,\cdots}^{\;l_1\quad\;\; 1\quad\;\; 2\quad \cdots\;\; l_1-1\;;\;\cdots}$$

$$=\varepsilon^{v(1,1)\,v(1,2)\cdots v(1,l_1)\,;\,\cdots}_{\;l_1\qquad 1\quad\;\cdots\;\; l_1-1\;;\;\cdots}$$

と書き直せるので，$\varepsilon(M_1)\boldsymbol{\cdot}\varepsilon(M_2)$ は

$$\varepsilon(M_1)\,\varepsilon(M_2)=\varepsilon^{1\;2\;3\;\cdots\;l_1-1\;\;l_1\;\;;\;\cdots}_{l_1\;1\;2\;\cdots\;l_1-2\;\;l_1-1\;;\;\cdots}=(-1)^{l_1-1}\boldsymbol{\cdot}(-1)^{l_2-1}\boldsymbol{\cdot}\cdots$$

$$=(-1)^{t}=\varepsilon(\mathrm{DACC}) \tag{A.5.5}$$

(l_k はみな偶数，そして $\sum_k l_k$ も偶数であることに注意)となる．注意すべきは ε(DACC)が互いに素な交互サイクルの個数 t だけで定まっていることである．

以上を総合すると，

$$\mathrm{Pf}\,A\boldsymbol{\cdot}\mathrm{Pf}\,B=\sum_{\substack{G_1\uplus G_2\text{ 上の}\\ \text{すべての DACC}}}\varepsilon(\mathrm{DACC})\,w(\mathrm{DACC}) \tag{A.5.6}$$

が得られる．

特別な場合を二つ挙げよう．

第一は $A=B$ の場合である．このとき

$$(\mathrm{Pf}\,A)^2=\sum_{\substack{G_1\uplus G_2\text{ 上の}\\ \text{すべての DACC}}}\varepsilon(\mathrm{DACC})\,w(\mathrm{DACC}) \tag{A.5.7}$$

であるが，$G=G_1\uplus G_2$ 上の偶数長のサイクル(交互とは敢えていわない)はそのまま交互サイクルとみなせる．(より厳密にいえば，G 上の交互サイクルはもちろん偶数長である；逆に G 上の偶数長のサイクルは——いつでも先頭の辺がG_1(すなわち E_1)に属し，他の辺は交互に $G_2, G_1, G_2, \cdots$ に属すと約束しておけば——交互サイクルに対応づけられる．) 結局(A.5.7)は，

$$(\mathrm{Pf}\,A)^2=\sum_{\substack{G\text{ 上の偶数長}\\ \text{のサイクルによる DCC}}}\varepsilon(\mathrm{DCC})\,w(\mathrm{DCC}) \tag{A.5.7$'$}$$

と書くことができる．このとき DCC を構成するサイクルの並べ方および各サイクルの先頭の選び方については，曖昧さを避けるために(A.4.9)，(A.4.8)の約束に従うものとする．

さて，(A.5.7′)において DCC を構成するサイクルの中に奇数長のものが含ま

れていてもよいとしたら何が起こるであろうか．そのようなDCCを一つ取り上げて，それの構成要素の奇数長のサイクルの中で先頭の点の番号が一番大きいものをC_kとする．C_kだけの向きを反転させ他のサイクルはそのままにしておいたものも一つのDCCである(これをDCC′と記す)．このとき$\varepsilon(\mathrm{DCC}')=\varepsilon(\mathrm{DCC})$(DCCを構成するサイクルの中で偶数長のサイクルの個数は変化しない！)であるが，$w(\mathrm{DCC}')=-w(\mathrm{DCC})$である($a_{ji}=-a_{ij}$であり$C_k$の長さが奇数であるから)．そこで$\varepsilon(\mathrm{DCC}')w(\mathrm{DCC}')=-\varepsilon(\mathrm{DCC})w(\mathrm{DCC})$で，$\varepsilon(\mathrm{DCC})w(\mathrm{DCC})+\varepsilon(\mathrm{DCC}')w(\mathrm{DCC}')=0$となる．このような対応は，奇数長のサイクルを含むようなすべてのDCCの集まりに対して一対一になっているから，結局(A.5.7′)の右辺の和$\sum$はG上のすべてのDCCについて取ることにしても結果には影響しない．そのような和は(A.4.4)により$\det A$に等しい．このようにして

$$(\mathrm{Pf}\,A)^2=\sum_{\substack{G\text{ 上のすべて}\\ \text{の DCC}}}\varepsilon(\mathrm{DCC})\,w(\mathrm{DCC})=\det A \qquad \text{(A.5.8)}$$

という関係(2.9.2)の“組合せ論的証明”が完成したことになる．

第二の場合として，(A.5.6)においてBが特殊な形をした行列

$$B^\circ=\begin{bmatrix} \begin{matrix}0&1\\-1&0\end{matrix} & & & O \\ & \begin{matrix}0&1\\-1&0\end{matrix} & & \\ & & \ddots & \\ O & & & \begin{matrix}0&1\\-1&0\end{matrix}\end{bmatrix} \qquad \text{(A.5.9)}$$

である場合を考える．$B^\circ=[b^\circ_{ij}]$と記すと，B°の非零要素は

$$b^\circ_{12}(=-b^\circ_{21})=b^\circ_{34}(=-b^\circ_{43})=\cdots=b^\circ_{n-1,n}(=-b^\circ_{n,n-1})=1$$

のみであるから，明らかに

$$\mathrm{Pf}\,B^\circ=b^\circ_{12}\,b^\circ_{34}\cdots b^\circ_{n-1,n}=1 \qquad \text{(A.5.10)}$$

である．また，G_2の辺は$E_2=\{e_1,\cdots,e_{n/2}\}$ ($\partial_2^+e_k=2k-1$，$\partial_2^-e_k=2k$；$k=1,\cdots,n/2$)であるから，G_2上の完全マッチングは

$$M_2=E_2 \quad (=M^\circ\text{ と表すことにする}) \qquad \text{(A.5.11)}$$

のみである．そこで，$G_1\cup G_2$上の交互サイクルC（長さlは偶数)は

$$C=(v_1,v_2,\cdots,v_l) \qquad (v_1<v_j\ \ (j\neq1)),$$

$$v_{2k+1}=\begin{cases} v_{2k}+1 & (v_{2k}\text{ が奇数のとき}),\\ v_{2k}-1 & (v_{2k}\text{ が偶数のとき})\end{cases} \qquad \text{(A.5.12)}$$

(ただし，$v_{l+1}=v_1$ とみなす)という形になっていなければならない．このような交互サイクル C を **M° 交互サイクル**(M°-alternating cycle)と呼ぶことにする．その重み $w(C)$ は

$$w(C)=\pm a_{v_1v_2}a_{v_3v_4}\cdots a_{v_{l-1}v_l} \tag{A.5.13}$$

で，± は

$$\text{“}v_{2k}=\text{偶数”であるような } k \text{ の個数が偶数なら }+\text{，奇数なら}- \tag{A.5.14}$$

と定める．((A.5.14)は「“v_{2k+1}＝奇数”であるような k の個数が偶数なら＋，奇数ならー」と述べることもできる．$v_{l+1}=v_1$ と約束したことにも注意．)

このような M° 交互サイクルで互いに素なもの $C_1,\cdots,C_t$ ですべての点を覆う(すべての点を通る，含む)ものを“互いに素な M° 交互サイクルによる被覆”と呼び DM°ACC(disjoint M°-alternating cycle cover)と記す．すると

$$\mathrm{Pf}\,A=\sum_{\text{すべての DM°ACC}}\varepsilon(\mathrm{DM^\circ ACC})\,w(\mathrm{DM^\circ ACC}) \tag{A.5.15}$$

という表現に到達する．($\varepsilon(\mathrm{DM^\circ ACC})=(-1)^t$ である．)

(A.5.6)，(A.5.15)は，それぞれ Pf A•Pf B および Pf A の定義式とみなすこともできる．(これらは行列式の表現(A.4.4)になぞらえられる．)

これらをそのまま計算するには，行列式の表現(A.4.4)でもそうであったように，行列の次数 n の多項式オーダーの手間では済まない．そこで前節 A.4 節において行列式の計算のために用いたのと類似の手法を Pfaffian の計算に適用して，除算を用いずに Pfaffian の計算を多項式オーダーの手間で実行する算法を作ってみよう[1]．

(A.5.13)から出発して Pf A を計算しようというわけであるが，基本的な着想は A.4 節とほとんど同じである．

DM°ACC の D(互いに素)と CC(サイクルによる被覆)の部分の条件を緩めて，単に M°AClWS(M°-alternating closed walk sequence)，すなわち M° の辺が一つおきに入っている有向閉路，すなわち M° 交互有向閉路を並べたものを

1)　主として Meena Mahajan, P. R. Subramanya and V. Vinay : A combinatorial algorithm for Pfaffians. *Electronic Colloquium on Combinatorial Complexity*, Report No. 30(1999), 22pp. および Günter Rote : Division-free algorithms for the determinant and the Pfaffian : Algebraic and combinatorial approaches. *Lecture Notes in Computer Sciences*, vol. 2122, pp. 119–135, Springer-Verlag, 2001 の一部による．

考える．

$$\text{M°ClWS}=\{C_1,\cdots,C_t\} \qquad (\text{A.5.16})$$

の $C_1,\cdots,C_t$ は交互有向閉路であるからそれぞれの長さは偶数である．これらに対しては(A.4.8)，(A.4.9)の条件と長さの総計が n に等しいという条件は同様に課しておく．M°AClWS(A.5.16)の重み w(M°AClWS)はそれを構成する C_i の重みの積，C_i の重みは(A.5.13)と同じ形で定義される．符号 ε(M°AClWS)は $(-1)^t$ であることも A.4 節におけると同様．

このとき(A.5.15)の値が

$$\text{Pf}'\,A=\sum_{\text{すべての M°ACIWS}}\varepsilon(\text{M°AClWS})\,w(\text{M°AClWS}) \qquad (\text{A.5.17})$$

に等しいことが，(A.4.4)と(A.4.12)が等しいことと“ほとんど同じ”ように証明できる．“ほとんど”という意味は，$M°$ 交互有向閉路 C_k が $M°$ 交互サイクルでない場合，すなわち A.4 節の(A.4.12)の後での解析の中の場合分け(a)，(b)において，(b)で発見される部分サイクルの長さが奇数であったときの扱いについてである．このときには，このサイクルを切り離して独立させることができない(奇数長では交互サイクルにはなりえないから)．しかしこの場合には C_k の中のこの部分サイクルの部分だけ向きを反転させたもう一つの $M°$ 交互有向閉路 C'_k が存在するはずで，M°AClWS の中の C_k を C'_k でおきかえた M°AClWS′ を作ることができる．この対応は一意に定まり，$w(C'_k)=-w(C_k)$ であるから w(M°AClWS′)$=-w$(M°AClWS)であり，かつ ε(M°AClWS′)$=\varepsilon$(M°AClWS)$(=(-1)^t)$であるので，両者は(A.5.17)の右辺の中で消し合うことになる．

そこで結局 Pf′ A=Pf A となることが，行列式の場合と同様に示されたことになる．

具体的な計算方法も観念的には同様で，DAG を構成しながら計算を進めて行けば Pf′ A の計算が完了する．ただ偶数番目の層と奇数番目の層とでは扱いが異なることに注意しなければならない．

行列 A の疎性がこの場合にも活用できること，計算量は一般には $O(n^4)$ であることなども，行列式の場合と同様である．

注意 A.5.1 (A.5.6)に対してもまったく同様の手順で計算できることも明らかであろう．行列 B の要素の値が ±1 と 0 のみでないことが異なるだけである．

索　　引

欧　文

ア 行

カ　行

タ　行

ナ　行

ハ　行

マ　行

ヤ　行

ラ　行

ワ　行

著者略歴

伊理正夫（いりまさお）

1933年　茨城県に生まれる
1960年　東京大学大学院数物系研究科応用物理学専門課程博士課程修了，
　　　　工学博士
現　在　東京大学名誉教授

〔主な著書〕
『Network Flow, Transportation and Scheduling—Theory and Algorithms』（Academic Press, 1969）
『数値計算—方程式の解法—』（朝倉書店，1981）
ほか

基礎数理講座 3
線形代数汎論　　　　定価はカバーに表示

2009年8月25日　初版第1刷

著　者　伊　理　正　夫
発行者　朝　倉　邦　造
発行所　株式会社　朝　倉　書　店
東京都新宿区新小川町 6-29
郵便番号　162-8707
電　話　03(3260)0141
FAX　03(3260)0180
http://www.asakura.co.jp

〈検印省略〉

新日本印刷・渡辺製本

ISBN 978-4-254-11778-3　C 3341　　Printed in Japan

前政策研究大学院大 刀根　薫著
基礎数理講座1
数　理　計　画
11776-9 C3341　A５判 248頁 本体4300円

理論と算法の緊密な関係につき，問題の特徴，問題の構造，構造に基づく算法，算法を用いた解の実行，といった流れで平易に解説。〔内容〕線形計画法／凸多面体と線形計画法／ネットワーク計画法／非線形計画法／組合せ計画法／包絡分析法

東工大 高橋幸雄著
基礎数理講座２
確　率　論
11777-6 C3341　A５判 288頁 本体3600円

難解な確率の基本を，定義・定理を明解にし，例題および演習問題を多用し実践的に学べる教科書〔内容〕組合せ確率／離散確率空間／確率の公理と確率空間／独立確率変数と大数の法則／中心極限定理／確率過程／離散時間マルコフ連鎖／他

前慶大 柳井　浩著
基礎数理講座４
数　理　モ　デ　ル
11779-0 C3341　A５判 224頁 本体3900円

物事をはっきりと合理的に考えてゆくにはモデル化が必要である。本書は，多様な分野を扱い，例題および図を豊富に用い，個々のモデル作りに多くのヒントを与えるものである。〔内容〕相平面／三角座標／累積図／漸化過程／直線座標／付録

阪大 和田昌昭著
現代基礎数学３
線　形　代　数　の　基　礎
11753-0 C3341　A５判 168頁 本体2800円

線形代数の基礎的内容を，計算と理論の両面からやさしく解説した教科書。独習用としても配慮。〔内容〕連立1次方程式と掃き出し法／行列／行列式／ユークリッド空間／ベクトル空間と線形写像の一般論／線形写像の行列表示と標準化／付録

学習院大 飯高　茂著
講座　数学の考え方３
線　形　代　数 基礎と応用
11583-3 C3341　A５判 256頁 本体3400円

２次の行列と行列式の丁寧な説明から始めて，３次，n次とレベルが上がるたびに説明を繰り返すスパイラル方式を採り，抽象ベクトル空間に至る一般論を学習者の心理を考えながら展開する。理解を深めるため興味深い応用例を多数取り上げた

前東工大 志賀浩二著
数学30講シリーズ２
線　形　代　数　30　講
11477-5 C3341　A５判 216頁 本体3600円

〔内容〕ツル・カメ算と連立方程式／方程式，関数，写像／２次元の数ベクトル空間／線形写像と行列／ベクトル空間／基底と次元／正則行列と基底変換／正則行列と基本行列／行列式の性質／基底変換から固有値問題へ／固有値と固有ベクトル／他

元東海大 草場公邦著
すうがくぶっくす２
線　型　代　数（増補版）
11462-1 C3341　A５変判 180頁 本体2700円

１「なぜ必要か」「どうしてこのようなことを考えるか」，２図形的，感覚的なイメージが伝わるよう，の２点に重点を置き執筆。〔内容〕行列式の話／線型空間の話／線型写像と行列／線型写像とその行列の標準形／計量空間とユニタリー行列

東海大 伊藤雄二著
新数学講座１
微　分　積　分　学
11431-7 C3341　A５判 312頁 本体4500円

微分積分学の論理的構成についてできるだけ正確かつわかりやすく解説されている。また計算機の発展に伴う数値計算の重要性を認識して数値計算，近似法について充分説明がなされ，線形代数とのつながりについても留意して解説されている

奥川光太郎著
基礎数学シリーズ７
線　形　代　数　学　入　門（復刊）
11707-3 C3341　B５判 214頁 本体3200円

直線・曲線・曲面など平面・空間でのテーマや応用例を豊富に取りあげ，線形代数の考え方を解説〔内容〕ベクトル／行列／行列式／行列式の積／行列の階数／座標変換／２次形式（ユニタリー空間／エルミート形式他）／付録（置換／斉次座標）他

神奈川大 小国　力・神奈川大 小割健一著
MATLAB数式処理による数学基礎
11101-9 C3041　A５判 192頁 本体3400円

数学・数式処理・数値計算を関連づけ，コンピュータを用いた応用にまで踏み込んだ入門書。〔内容〕微積分の初歩／線形代数等の初歩／微積分の基礎／積分とその応用／偏微分とその応用／3変数の場合／微分方程式／線形計算と確率統計計算

上記価格（税別）は 2009 年 7 月現在